U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

MSc

2.º
CICLO

FCUP
IPATIMUP
2014

U.PORTO

Contribution to the mutational spectrum of Maple Syrup Urine Disease and functional characterization of the alteration c.108+6T>C in *BCKDHA* gene.

Ana Luísa Archer Carvalho Castro Taveira

FC

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

25 ANOS
A DESCOBRIR O CANCRO
ipatimup

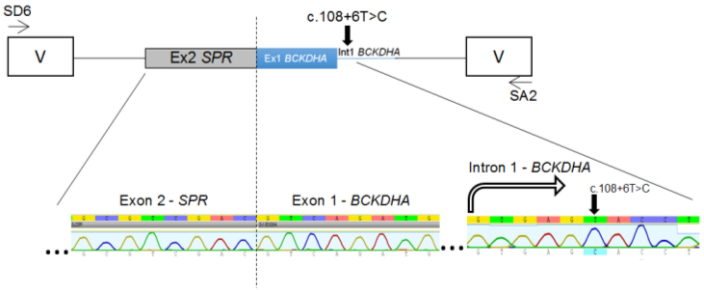
Contribution to the mutational spectrum of Maple Syrup Urine Disease and functional characterization of the alteration c.108+6T>C in *BCKDHA* gene.

Ana Luísa Archer Carvalho Castro Taveira

Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto em
Biologia Celular e Molecular

2014

MSc



Contribution to the mutational spectrum of Maple Syrup Urine Disease and functional characterization of the alteration c.108+6T>C in *BCKDHA* gene.

Ana Luísa Archer Carvalho Castro Taveira
Mestrado em Biologia Celular e Molecular
Departamento de Biologia
2014

Orientador
Doutora Maria Sofia Pacheco Quental, Investigadora Pós-Douturamento, Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP)

Coorientador
Professora Doutora Maria João Prata Martins Ribeiro, Professora Associada, Faculdade de Ciências da Universidade do Porto e Investigadora, Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP)

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Dissertação para a candidatura ao grau de Mestre em Biologia Celular e Molecular submetida à Faculdade de Ciências da Universidade do Porto.

O presente trabalho foi desenvolvido sob a orientação científica da Doutora Maria Sofia Ferreira Pacheco Quental e co-orientação da Prof. Doutora Maria João Prata Martins Ribeiro e foi realizado no Instituto de Patologia e Imunologia Molecular da Universidade do Porto.

Dissertation to apply for the Master's Degree in Cell and Molecular Biology, submitted to the Faculty of Sciences of the University of Porto.

The present work was developed under the scientific Supervision of Doctor Maria Sofia Ferreira Pacheco Quental and co-supervision of Prof. Doctor Maria João Prata Martins Ribeiro and was done at the Institute of Molecular Pathology and Immunology of University of Porto.

“Que os vossos esforços desafiem as impossibilidades,
lembrai-vos de que as grandes coisas do homem
foram conquistadas do que parecia impossível.”

Charles Chaplin

Agradecimentos

À Sofia, minha orientadora, por ter aceitado orientar uma das etapas mais importantes do meu percurso académico e por ter sido uma pessoa incansável nesta etapa. Obrigada por toda a preocupação, todo o empenho, todos os ensinamentos, toda a orientação, todo o apoio, toda a paciência, todas as sugestões e por seres a ótima pessoa que és. Nunca me senti sozinha nesta luta, pois sabia que podia sempre contar contigo. Obrigada! A ciência precisa de pessoas como tu, que ajudam os outros com toda a vontade e entusiasmo.

À Prof. Doutora Maria João Prata, por ter aceitado ser minha co-orientadora e por estar sempre disponível para o que precisasse. Obrigada por todos os ensinamentos e por todas as sugestões feitas durante a realização desta dissertação de mestrado e por toda a simpatia com que sempre me recebeu.

À Liliana por toda a ajuda no início do trabalho e ao longo deste sempre que era necessário.

Ao Prof. Doutor José Pissarra, diretor do Mestrado em Biologia Celular e Molecular da Faculdade de Ciências da Universidade do Porto, por todo o trabalho realizado em prol deste mestrado.

Ao corpo docente do Mestrado em Biologia Celular e Molecular por todos os ensinamentos dados que me foram úteis durante a realização desta dissertação de mestrado e que o serão com certeza para o futuro.

Ao grupo de Genética Populacional do IPATIMUP, por me terem recebido e por estarem sempre prontos a ajudar e obrigada também a alguns membros de outros grupos que ajudaram sempre que solicitado.

Aos meus Pais, Eduardo e Glória, e à minha irmã Sara por todo o apoio, não só ao longo desta etapa, mas ao longo de toda a minha vida. Obrigada por construírem um lar onde me sinto sempre segura e feliz!

A toda a restante família por todo o apoio e por terem ajudado a formar a pessoa que sou hoje.

Aos meus amigos por todo o apoio, por todos os momentos de amizade bem passados que fazem a vida valer a pena, por todas as gargalhadas dadas e por todos os momentos de incentivo.

Resumo

A leucinose é uma doença autossómica recessiva que envolve o metabolismo dos aminoácidos ramificados (BCAAs) – leucina, isoleucina e valina – e cuja incidência na maioria das populações é de 1/185.000 recém-nascidos. Mutações num dos três genes- *BCKDHA*, *BCKDHB* e *DBT* – que codificam os componentes catalíticos E1α, E1β e E2, respetivamente, do complexo da desidrogenase dos α-cetoácidos ramificados (BCKD) podem causar leucinose.

Neste trabalho, fez-se a caracterização molecular de 2 doentes recentemente diagnosticados com leucinose. Num deles, de origem angolana, foi detetada uma deleção de 11 bp (c.93_103del11; p.Ala32Phefs*48) em homozigotia no gene *BCKDHB*, que já tinha sido anteriormente descrita. No outro paciente, de origem portuguesa, identificaram-se duas mutações em heterozigotia, c.799C>T (p.Gln267X) e c.359T>C (p.Phe120Ser), também no gene *BCKDHB*. Como a mutação c.359T>C nunca tinha sido descrita, para avaliar o seu possível efeito patogénico, usaram-se vários programas bioinformáticos e analisou-se o efeito da alteração na estrutura 3D da proteína, tendo sido obtido resultados consistentemente apontando para o efeito causal da mutação.

Por outro lado, foi efetuada a análise funcional da mutação c.108+6T>C do gene *BCKDHA*. Esta alteração tinha sido detetada pela primeira vez num doente de origem brasileira também diagnosticado com leucinose. Tendo as análises preliminares *in silico* indicado que a alteração deveria muito provavelmente alterar o normal padrão de *splicing* do gene em que ocorria, no sentido de confirmar o efeito patogénico recorreu-se à construção de minigenes através da clonagem do exão 1 e parte do intrão 1 do gene *BCKDHA*, com e sem mutação, no vetor pSPL3. O estudo do minigene em células HeLA e AGS e posterior extração do RNA e síntese do cDNA possibilitou a análise dos transcritos processados. Pela comparação dos transcritos produzidos pelo minigene WT e pelo minigene com a mutação, concluiu-se que a mutação comprometia o processo de *splicing* da zona genómica em estudo, causando o *skipping* do exão clonado no vector pSPL3. Deste modo, afigura-se altamente expectável que *in vivo* a mutação c.108+6T>C faça com que o local de *splicing* onde se localiza deixe de ser reconhecido, causando o *skipping* do exão 1.

Este trabalho veio enriquecer o conhecimento sobre o espectro mutacional da leucinose, uma vez que resultou na descrição de duas novas mutações, uma delas *missense* (c.359T>C; p.Phe120Ser em *BCKDHB*) e a outra de *splicing* (c.108+6T>C em *BCKDHA*). Para a última, foi

ainda possível confirmar a nível funcional o seu efeito ao nível do mRNA do gene *BCKDHA,* contribuindo para reforçar o papel da desregulação do *splicing* como agente responsável por doenças.

O trabalho permitiu também demonstrar a eficácia da estratégia baseada na construção e clonagem de minigenes para avaliar o efeito de mutações suspeitas de afetar o *splicing.* O esforço agora desenvolvido para estabelecer adequadamente os procedimentos técnicos necessários, permitirá no futuro a utilização dos minigenes para investigar o efeito de outras mutações candidatas a interferir no processo de *splicing*.

Palavras-chave: *BCKDHA*; *BCKDHB*; Leucinose; Minigene; Mutação de splicing.

Abstract

Maple syrup urine disease (MSUD) is an autosomal recessive disease that affects the metabolism of branched-chain amino acids (BCAAs) – leucine, isoleucine and valine, whose incidence is about 1/185.000 live newborns in most populations. Mutations in any of the three different genes – *BCKDHA*, *BCKDHB* and *DBT* – encoding respectively for the E1α, E1β and E2 catalytic components of the branched-chain α-ketoacid dehydrogenase (BCKD) complex can cause maple syrup urine disease (MSUD).

In this study, the molecular characterization of two newly diagnosed patients with MSUD was conducted. In one of them, with Angolan origin, a previously described deletion of 11 bp (c.93_103del11; p.Ala32Phefs*48) was identified in homozygosity in *BCKDHB* gene. In the other patient, with Portuguese origin, two mutations in heterozygosity were identified also in *BCKDHB*, c.799C>T (p.Gln267X) and c.359T>C (p.Phe120Ser). The c.359T>C mutation had never been described and as so, to assess their possible pathogenic effect several bioinformatics programs were used as well as it was performed 3D structure analysis, obtaining results consistently pinpointing the disease causative effect of the mutation.

Furthermore, the functional analysis of the substitution c.108+6T>C in intron 1 of *BCKDHA* previously identified in a patient from Brazil diagnosed with MSUD was also performed. Since preliminary *in silico* analyses had indicated that it probably induced an alteration in the normal splicing pattern of the gene, in order to confirm the previsions, a minigene system was constructed by cloning exon 1 and part of intron 1 of *BCKDHA* gene, with and without the substitution, in pSPL3 vector.

The study of this minigene in HeLa and AGS cells and subsequent RNA extraction and cDNA synthesis allowed the analysis of the processed transcripts. Through the comparison of the transcripts produced by the WT minigene and by the mutant one, it was possible to conclude that the alteration indeed affected the splicing process in the gene since the presence of the c.108+6T>C mutation resulted in the skipping of the exon cloned into pSPL3 vector. Therefore, it seems very likely that *in vivo* the mutation causes the skipping of exon 1 of *BCKDHA* gene.

This work contributed to increase the knowledge on the mutational spectrum of MSUD since it led to identify two new alterations, one missense mutation (c.359T>C; p.Phe120Ser in *BCKDHB*) and a splicing mutation (c.108+6T>C in *BCKDHA*). Concerning this latter, its pathogenic impact was in addition functionally demonstrated.

Furthermore, this work has evidenced the efficiency of a strategy based on the construction of minigene systems to evaluate the effect of a mutation of uncertain pathogenic impact on splicing. The efforts invested in the implementation of reliable methodologies, will allow in the future to extend the approach to other mutations that might interfere with splicing.

Key-words: *BCKDHA*; *BCKDHB;* Maple Syrup Urine Disease; Minigene; Splicing mutation.

Table of contents

Table Index

Figure Index

Abbreviations

®	registered trademark
µl	microliter
ml	milliliter
aa	amino acids
ABCA1	ATP-binding cassette sub-Family A, member 1
AGS	adrenogenital syndrome
ALDH7A1	aldehyde dehydrogenase 7 family, member A1
Arg	arginine
AS	alternative splicing
ATP	adenosine triphosphate
BCAAs	branched chain amino acids
BCATs	branched chain aminotransferase isozymes
BCKAs	branched chain α-ketoacids
BCKD	branched chain α-ketoacid dehydrogenase
BCKDHA	branched chain keto acid dehydrogenase E1, alpha polypeptide
BCKDHB	branched chain keto acid dehydrogenase E1, beta polypeptide
BCKDK	branched chain ketoacid dehydrogenase kinase
BDGP	Berkeley drosophila genome project
bp	base pair
°C	degree Celsius
cDNA	complementary deoxyribonucleic acid
DBT	dihydrolipoamide branched chain transacylase
DMEM	Dulbecco's Modified Eagle Medium
ddNTPs	dideoxynucleotide triphosphates

DLD	dihydrolipoamide dehydrogenase
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
EDTA	ethylenediamine tetraacetic acid
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
et al	et alii
ExoSAP	Exonuclease I and Shrimp Alkaline Phosphatase
FAD	flavin adenine dinucleotide
FBS	fetal bovine serum
FCUP	Faculdade de Ciências da Universidade do Porto
Glu	glutamine
HeLa	Henrietta Lacks
hnRNP	heterogeneous nuclear ribonucleoprotein
HSF	human splicing finder
Ile	isoleucine
IPATIMUP	Institute of Molecular Pathology and Immunology at the University of Porto
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
KIC	α-ketoisocaproate
KIV	α-ketoisovalerate
KMV	α-keto-β-methylvalerate
LB	lysogeny broth
LDLR	low density lipoprotein receptor
Leu	leucine
Lys	lysine

NaCl	sodium chloride
NMD	nonsense-mediated decay
mRNA	messenger ribonucleic acid
MS/MS	tandem mass spectrometry
MSUD	maple syrup urine disease
NAD	nicotinamide adenine dinucleotide
NNSplice	splice site prediction by neural network
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDC	pyruvate dehydrogenase complex
Phe	phenylalanine
PP2Cm	protein phosphatase 2Cm
PPM1K	protein phosphatase 1K
RNA	ribonucleic acid
rpm	rotations per minute
RPMI	Roswell Park Memorial
RT	reverse transcriptase
RT-PCR	reverse transcription polymerase chain reaction
SDM	site directed mutater
Ser	serine
snRNA	small nuclear ribonucleic protein
snRNP	small nuclear ribonucleoproteins
SOC	super optimal broth with catabolic repressor
SPR	sepiapterin reductase
ss	splice site
TCA	tricarboxylic acid
Val	valine

WT	wild type

1. Introduction

1.1. Branched-chain amino acids metabolism

The branched-chain amino acids (BCAAs), leucine, isoleucine and valine, are neutral, aliphatic amino acids with a branched methyl group in the side chain.

BCAAs are essential, also called indispensable, amino acids that need to be obtained from the diet in sufficient amounts to promote normal growth and development. Usually they comprise about 15%-20% of the total protein intake from food (Layman, 2003). BCAAs play important roles in cellular function, especially leucine since it promotes protein synthesis, inhibits protein degradation and stimulates insulin release (Mordier *et al.*, 2000; Proud, 2002; Lynch *et al.*, 2002).

The first step in the catabolism of leucine, isoleucine and valine (Figure 1) is a reversible transamination catalyzed by branched-chain aminotransferase isozymes (BCATs), including a mitochondrial form with ubiquitous distribution and a cytosolic BCAT present mainly in the nervous system, to form their respective branched-chain α-ketoacids (BCKAs): α-ketoisocaproate (KIC), α-keto-β-methylvalerate (KMV) and α-ketoisovalerate (KIV) (Harper *et al.*, 1984). The cytosolic BCKAs are then transported across the mitochondrial membrane (Hutson and Hall, 1993). Once in the mitochondria, branched-chain α-ketoacid dehydrogenase (BCKD) complex catalyzes the oxidative decarboxylation of BCKAs, generating their respective branched-chain acyl-CoAs, that are further metabolized via separate pathways (Figure 1) (Harris *et al.*, 1997).

The step catalyzed by BCKD commits the BCAAs to degradation, being so the rate-limiting step in the BCAA catabolism (Harris *et al.,* 1990) (Chuang and Shih, 2001). The BCKD complex is therefore the most important regulatory enzyme in the catabolic pathway of the BCAAs (Harris *et al.*, 2005).

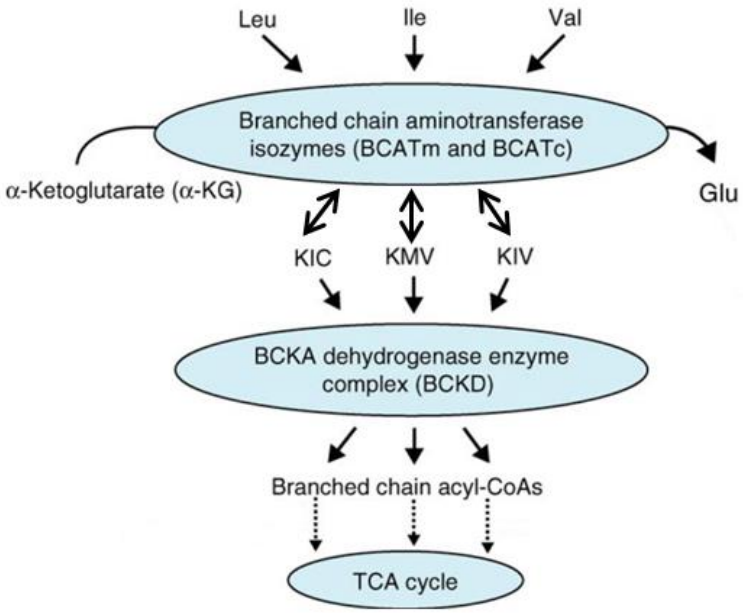


Figure 1. BCAA catabolic pathways. The first step is transamination of the BCAAs (leucine (Leu), isoleucine (Ile), valine (Val) to produce the branched-chain α-keto acids (BCKAs): α-ketoisocaproate (KIC), α-keto-β-methylvalerate (KMV) and α-ketoisovalerate (KIV). α-ketoglutarate is the α-keto acid acceptor of the BCAA α-amino group and glutamate (Glu) is the product. The second catabolic step is oxidative decarboxylation, catalyzed by the branched chain α-keto acid dehydrogenase (BCKD) complex. Leucine is ketogenic, forming acetyl-CoA and acetoacetate, valine is glucogenic, entering the tricarboxylic acid (TCA) cycle as succinyl-CoA, whereas isoleucine is both ketogenic and glucogenic. Adapted by Hutson *et al.*, 2005.

1.2. Branched-chain α-ketoacid dehydrogenase complex

The human branched-chain α-ketoacid dehydrogenase (BCKD) catalytic machine is a member of the highly conserved mitochondrial α-ketoacid dehydrogenase complexes comprising BCKD complex, pyruvate dehydrogenase complex (PDC) and α-ketoglutarate dehydrogenase complex (Reed *et al.*, 1985).

The BCKD complex is composed by three catalytic components - a decarboxylase (E1) that has a heterotetrameric structure formed by two E1α and two E1β subunits (α2β2), a dihydrolipoyl transacylase (E2) and a dihydrolipoamide dehydrogenase (E3). The complex has also two regulatory enzymes, which are a specific kinase (BCKDK) and a specific phosphatase (PP2Cm) (Yeaman, 1989) (Figure 2).

Components E1, E2 and the regulatory enzymes are specific of BCKD, while E3 is shared by the three α-ketoacid dehydrogenase complexes and by the glycine cleavage complex (Yeaman, 1989; Danner and Doering 1998).

The complex is organized around a 24-meric cubic core of E2, to which 12 copies of E1, 6 copies of E3, and the kinase and the phosphatase are attached (Yeaman, 1989; Danner and Doering 1998).

BCKD complex requires several cofactors to function properly, including thiamine pyrophosphate (for the E1 component), coenzyme A (for the E2 component), lipoamide, flavin and nicotinamide adenine dinucleotides (FAD and NAD) for the E3 component.

Activity of the complex is controlled by covalent modification through phosphorylation of a serine residue (Ser-337 previously identified as Ser-292) of the E1α subunit performed by the specific kinase of the complex that causes inactivation of the E1 component and consequently of the complex; dephosphorylation by the specific phosphatase leads to the reactivation of BCKD (Damuni *et al.*, 1984; Popov *et al.*, 1992). The extent to which the complex is dephosphorylated and therefore catalytically active depends on the relative activities of PP2Cm and BCKDK (Harris *et al.*, 2001). At low BCAA levels, E1α is hyperphosphorylated by BCKD kinase, leading to lower BCKD activity and reduced loss of BCAA. At elevated BCAA levels, E1α is dephosphorylated by BCKD phosphatase, to induce BCKD complex activity and the removal of excess of BCAA. Therefore, BCKD phosphorylation/dephosphorylation is critical to BCAA homeostasis (Sun *et al.*, 2011).

All components of BCKD complex are encoded by nuclear genes. The human E1α gene (*BCKDHA*) (MIM #608348) is located at chromosome 19 (19q13.1-13.2) and has 9 exons leading to an mRNA that is translated into a protein with 445 residues (Fekete *et al.*, 1989). The *BCKDHB* gene (MIM #24861), which encodes the E1β protein, is located at the chromosome 6 (6q14.1) (Zneimer *et al.*, 1991), has 11 exons and originates a protein with 392 residues (Edelmann *et al.,* 2001). The E2 protein is codified by the *DBT* gene (MIM #248610) that is located at the chromosome 1 (1p31). This gene is composed by 11 exons and originates a protein with 482 aa (Zneimer *et al.*, 1991). *DLD* (MIM #238331) at chromosome 7 (7q31-32) encodes the E3 protein, and contains 14 exons forming a protein with 509 aa (Scherer *et al.,* 1991).

Concerning the regulatory subunits of BCKD complex, the gene for the specific BCKD kinase (*BCKDK*) is located on chromosome 16 at 16p11.2 and has 13 exons encoding a protein of 412 residues (Suryawan *et al.,* 1998). The gene for the BCKD specific phosphatase (*PPM1K*) is

located on chromosome 4 and consists of 7 exons that are translated into a protein with 372 residues (Damuni *et al*., 1984) (Lu *et al.,* 2009).

Genetic defects leading to impairment of the BCKD complex are responsible for a condition known as Maple Syrup Urine Disease (MSUD) (Harris *et al.,* 2004).

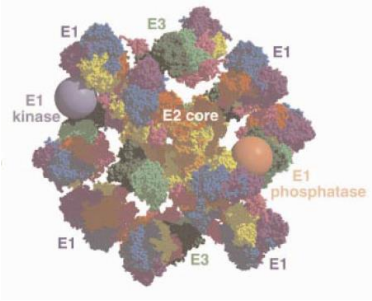


Figure 2. Representation of the branched chain α-ketoacid dehydrogenase (BCKD) complex. Adapted from (Ævarsson *et al.,* 2000).

1.3. Maple Syrup Urine Disease

Inherited metabolic diseases are caused by the deficiency or even absence of activity of a particular enzyme in a metabolic pathway. The blockage can lead to the accumulation of toxic substances in body fluids and tissues or lack of essential metabolites, many times leading to implications for the mental and/or physical development of affected individuals. The majority of these disorders is characterized by an autosomal recessive pattern of inheritance.

Among them is maple syrup urine disease (MSUD; OMIM 248600), an autosomal recessive disorder, first described in 1954 by Menkes and collaborators (Menkes *et al.*, 1954) that is associated to dysfunction of branched-chain amino acids metabolism. The characteristic odor of the urine of the patients, resembling maple syrup, gave the name to the disease (Dancis *et al.,* 1959). The metabolic block causing MSUD was then identified to occur in the decarboxylation of BCKAs, due to the deficient activity of BCKD complex (Dancis *et al.,* 1960) (Menkes, 1959).

1.3.1. Clinical phenotypes

The classic form of MSUD is the most severe subtype with neonatal onset. Affected newborns have 0-2% of normal BCKD complex activity (Packman *et al.*, 2012) and although

apparently healthy at birth the progressive accumulation of BCKAs and BCAAs conduct to the development of the symptoms in the first days of life. Patients present poor feeding and irritability that, if untreated, may progress to lethargy, coma and ultimately death. This subtype of MSUD has the highest morbidity and is the most common, accounting for about 75% of the patients (McLaughlin *et al.*, 2013).

Patients with the intermediate form have consistently abnormal BCAAs profile and neurologic deterioration, but do not develop catastrophic illness in the initial days of life as observed in classic MSUD. BCKD complex activity is usually higher than in the previous form, lying between 3 to 30% of normal. Most patients are diagnosed between five months and seven years of age (Chuang and Shih, 2001).

Intermittent MSUD patients show normal early development, growth and intelligence but metabolic decompensation can occur as a result of any physiologic stress. In spite of that, when asymptomatic these individuals usually tolerate a regular diet and have normal levels of plasma BCAAs. Activity of the BCKD complex is reportedly 5 to 20% of normal, with the diagnosis usually being established between 5 months and 2 years of age. During acute crises, behavioral changes may progress to seizures and coma and at this time the biochemical profile is similar to that of classic MSUD (Chuang and Shih 2001).

The thiamine-responsive form was first described by Scriver and collaborators (1971) in patients in whom the hyperaminoacidemia was completely corrected by administration of thiamine hydrochloride. As occurs in the intermediate type there is no neonatal illness and BCKD complex activity is about 30 to 40% of the normal rate (Chuang *et al.*, 1982). Patients with this phenotype are usually treated with combined therapy of dietary adjusted restriction and thiamine supplementation.

The E3-deficient type is associated with combined enzyme deficiencies in pyruvate dehydrogenase, α-ketoglutarate dehydrogenase and BCKD complexes because E3 is a common component among these mitochondrial α-ketoacid dehydrogenase complexes (Yeaman, 1989). Patients with this disease form present a severe phenotype distinct from MSUD that is characterized by lactic acidosis and progressive neurologic deterioration (Robinson *et al.,* 1977).

1.3.2. Diagnosis

In newborns or young children manifesting clinical symptoms suggestive of MSUD, diagnosis confirmation is performed through amino acid and organic acid analysis profile, by measuring the plasma BCAAs concentrations that are usually elevated five to ten-fold greater than normal and by detection of alloisoleucine that has been considered as a pathognomonic marker for the confirmation of the diagnosis (Snyderman *et al.,* 1964; Strauss and Morton, 2003; Chuang *et al.,* 2006). A general disturbance of amino acid concentration ratios is also observed (Chuang and Shih, 2001, Morton *et al.,* 2002) and high levels of the BCKAs derived from the BCAAs can also be detected in the urine.

However, in the last decades the identification or suspicion of MSUD patients become possible even before the manifestation of the first symptoms since the disease has been included in the expanded newborn screening programs already implemented in many countries, including Portugal (Vilarinho *et al.*, 2006). For the screening, it is used tandem mass spectrometry (MS/MS) whose results permits to identify newborn suspects who may develop MSUD in the first days of life, therefore allowing immediate therapeutic implementation. The inclusion of MSUD in the MS/MS newborn screening program represents a major improvement since early diagnosis and intervention during the pre-symptomatic or early symptomatic period improves the outcome of the patients and usually prevents irreversible brain damage (Schadewaldt *et al.*, 1999; Simon *et al.*, 2006).

In patients with positive results in the neonatal screening, the biochemical diagnosis (previously described) must be performed to confirm the diagnosis.

Whenever possible, the deficiency of the branched-chain α-ketoacid dehydrogenase enzyme complex activity may also be measured in cultured fibroblasts or lymphocytes as an additional method for diagnosis confirmation.

Molecular diagnostic allows the identification of causative mutations, which is crucial to provide prenatal diagnosis in high-risk pregnancies and carrier detection in family members to offer them genetic counseling.

1.3.3. Treatments

As for the majority of metabolic diseases, therapy for MSUD comprises long-term disease-specific approaches that consists mainly in dietary restriction of the amino acids that cannot be properly degraded, particularly of leucine, and replacement of deficient substrates, which in the

case of MSUD was firstly proposed by Snyderman (Snyderman *et al.,* 1964). Besides, emergency treatments are usually implemented during metabolic decompensation crisis.

Since the intellectual outcome of patients is dependent on the time that leucine levels remain above the normal, to protect the neonatal brain of the acutely ill newborn from permanent damage an emergency treatment is used to reduce leucine concentration as rapidly as possible. This intervention consists mainly in exogenous toxin removal, because a high-energy enteral or parenteral nutrition alone is insufficiently effective to rapidly lower plasma leucine levels (Berry *et al.,* 1991). For metabolic removal in acute situations, peritoneal dialysis or venous hemofiltration can be performed while high calorie fluids are provided intravenously to inhibit catabolism and promote anabolism.

Besides this emergency intervention, long term dietary therapy is also used in MSUD patients during all life to normalize the concentration of blood BCAAs and to provide adequate nutrition to promote normal growth and development, especially in young patients. This treatment consists fundamentally in a BCAA-restricted diet that should be adjusted to each patient individual requirements viewing which BCAAs levels must be regularly monitored. Thiamine supplementation should be used to test for thiamine-responsiveness of the patients. Several commercial synthetic formulas are already available for MSUD patients, which are enriched with the amino acids that compete with BCAA transport and help to maintain physiological amino acid plasma levels and transport into the brain (Strauss *et al.*, 2010).

In asymptomatic individuals with mild variant forms of MSUD the necessity of long-term treatment is yet unclear, but interventions such as prevention of metabolic crisis, careful instructions of affected families and evaluation to identify individuals at risk for metabolic decompensation are recommended (Vockley *et al.,* 2006; Simon *et al., 2006).*

Interestingly, liver transplantation has been suggested as an effective treatment for classic MSUD. This discovery was made from the observation of various patients with MSUD who required liver transplantation for liver failure and demonstrated a positive metabolic outcome (Strauss *et al.,* 2006; Diaz *et al.,* 2014). Since then, numerous liver transplants have been performed and the effectiveness of this procedure in MSUD suggests that providing 9-13% of the bodies' BCKD complex activity is sufficient to restore BCKD complex homeostasis (Burrage *et al.*, 2014). However, liver transplantation has important risks, such as postoperative complications and the need for long-term immunosuppression, which must be carefully evaluated and considered.

1.3.4. Incidence and mutational spectrum

MSUD is a rare disorder, affecting around 1/185.000 live births, thought it can attain very high incidence in populations or communities where endogamous practices are common (Chuang and Shih 2001). That is the case of some Old Order Mennonite communities, where a high incidence of classic MSUD is observed due to the alteration p.Tyr438Asn in E1α (Marshall and DiGeorge, 1981; Zhang *et al.,* 1989) or the p.Arg183Pro in E1β, affecting the Ashkenazy Jewish (Edelmann *et al*., 2001). In Portugal, MSUD has been estimated to affect 1/86.800 live newborns, which represents an incidence slightly higher than reported in most populations, but that can be mainly explained by the presence of a founder mutation (c.117delC in *BCKDHA*) in a Gypsy community from the South of the country (Quental *et al.,* 2008).

MSUD is mainly caused by recessively inherited mutations in any of the three genes that code for catalytic subunits specific of the BCKD complex namely *BCKDHA* (E1α), *BCKDHB* (E1β) and *DBT* (E2). Recently, however, a patient with a mild variant phenotype was reported to have a mutation in the BCKD specific phosphatase (PP2Cm, *PPM1K* gene) (Oyarzabal *et al.,* 2013).

A comprehensive knowledge of the MSUD mutational spectrum has important implications for disease management and genetic counseling (Narayanan *et al*., 2013).

Up to now, more than 150 pathogenic alterations underlying MSUD have been identified, distributed in a near equal number in the three genes that encode the catalytic subunits (Quental, 2009).

According to the data available in the Human Gene Mutation Database (version 2013.4; http://www.hgmd.org/), the majority of the alterations that cause MSUD comprises single nucleotide substitutions (~74%), including the commonly referred “splicing mutations”, which are those that disrupt the normal splicing or alternative splicing of gene’ transcripts. In the last years, much attention has been given to the mechanisms by which disruption of pre-mRNA splicing play a role in human disease, being increasingly acknowledged that the understanding of the pathobiology of splicing expands opportunities for therapeutic intervention by either directly addressing the cause or by providing novel approaches to circumvent disease processes (Ward and Cooper, 2010).

1.4. Splicing process

Expression of eukaryotic structural genes is a multistep process that includes transcription of the gene, splicing and polyadenylation of the primary transcript, and transport of the fully processed mRNA to the cytoplasm (Singer and Green, 1997).

Most genes in higher eukaryotes are transcribed as precursor mRNAs (pre-mRNAs) that contain intervening sequences (introns) as well as expressed sequences (exons). The process of pre-mRNA splicing, discovered in the late 1970s, briefly consists in a series of reactions whereby the intronic RNA segments are removed and discarded while the remaining exonic segments are joined end-to-end, to give a shorter RNA product (Strachan T. and Read A. (2010) Human Molecular Genetics, 4th edition) .The precision and complexity of intron removal during pre-mRNA splicing still amazes even several years after its discovery.

The spliceosome is the macromolecular machine that performs the two primary functions of splicing: recognition of the intron/exon boundaries and catalysis of the cut-and-paste reactions that remove introns and join exons (Faustino and Cooper, 2003). It is a dynamic ribonucleoprotein complex composed by five small nuclear ribonucleoproteins (snRNPs) (U1, U2, U4, U5 and U6), each of which contains an RNA known as a snRNA and an elevated number of other non-snRNP proteins (Matera and Wang, 2014).

For an efficient splicing process, both *cis*-regulatory sequences and *trans*-acting factors are needed. The *cis* elements include highly conserved sequences such as donor and acceptor sites positioned at exon-intron boundaries, the branch point and the polypyrimidine tract, as well as auxiliary elements, that include exonic and intronic splicing enhancers (ESEs and ISEs, respectively) and silencers (ESS and ISSs, respectively) (Figure 3). *Trans*-acting elements may be divided into spliceosome and splicing factors which intervene through binding to splicing enhancers and silencers and include members of the well-characterized Ser/Arg-rich and heterogeneous nuclear ribonucleoprotein (hnRNP) protein families (Kornblihtt *et al.*, 2013).

The specificity of the splicing reaction is established by RNA-RNA base pairing between the RNA transcript to be spliced and snRNA molecules within the spliceosome.

First, U1 binds to the 5’ splice site, while U2 binds the intron branch site via RNA:RNA interactions between the snRNA and the pre-mRNA. Interaction between the splice donor and splice acceptor sites is stabilized by the binding of a multi-snRNP particle that contains the U4, U5 and U6 snRNAs. The U5 snRNP binds simultaneously to both the splice donor and splice acceptor sites. Their cleavage releases the intronic sequence and allows both exons to be spliced

together (Figure 4) (Faustino and Cooper, 2003); (Strachan T. and Read A. (2010) Human Molecular Genetics, 4th edition).

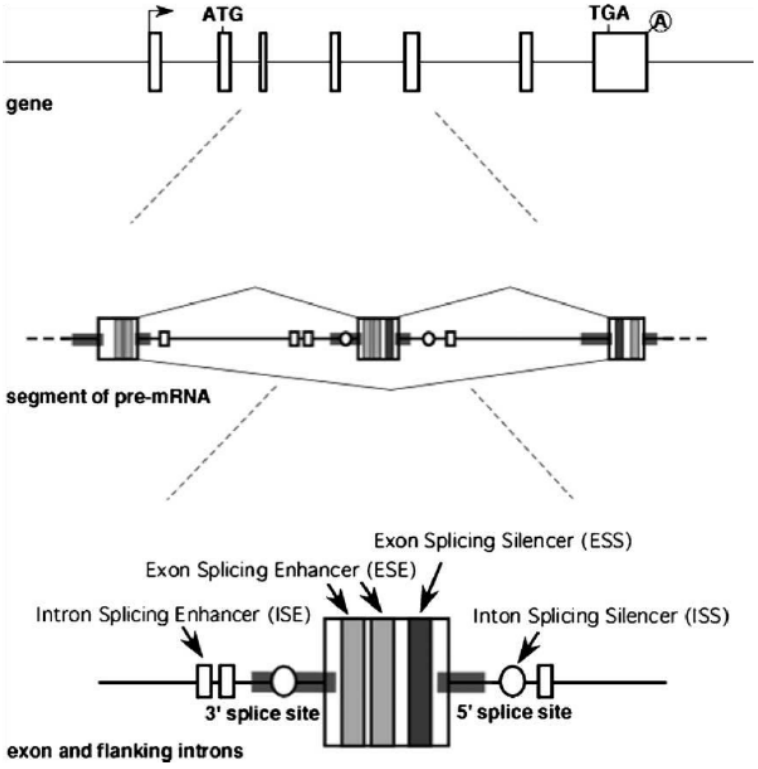


Figure 3. Exonic and intronic elements (Cooper, 2005)

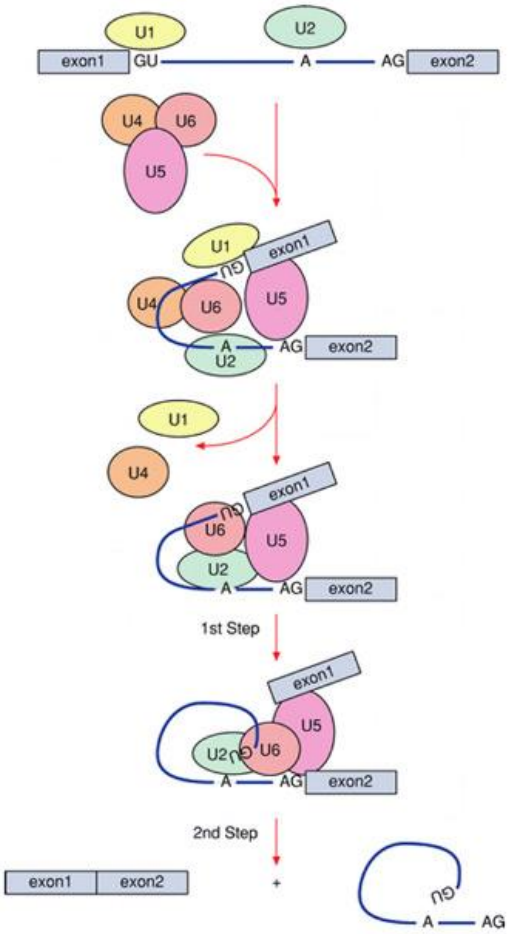


Figure 4. A simplified view of the spliceosome and process of splicing. Splicing involves several RNA-protein complexes, called small nuclear ribonucleoproteins (snRNPs), which together make up the spliceosome. U1 snRNP binds to the boundary between exon 1 and the intron by recognizing a specific sequence. U2 snRNP subsequently binds to the branch site (A) and then U4/U5/U6 triple snRNPs join in. After a dynamic rearrangement, U1 and U4 are destabilized, and the remaining snRNP complex is activated for the two steps that remove the intron and stitch together exons 1 and 2 (Gottlieb, 2003).

1.5. Alternative splicing

Through the process of alternative splicing (AS), different combinations of exons to use in the final mRNA transcript can be generated, creating different isoforms of a single gene (Douglas and Wood, 2011), and as so AS is considered one major source of protein diversity explaining the immense mammalian proteomic complexity. In fact, more than 90% of human genes undergo alternative splicing (Luco *et al.,* 2011).

Almost all protein-coding genes contain introns that are removed in the nucleus by RNA splicing during pre-mRNA processing but, as mentioned, exon usage is often alternative because a part of pre-mRNA can be removed as an intron or included in the mature mRNA as an alternative exon (Ben-Dov *et al.,* 2008). The relative positions of weak and strong splice sites give rise to the different modes of alternative splicing including: inclusion of alternative cassette exons that can be independently included or excluded from the mRNA (Figure 5.A); inclusion of alternative mutually exclusive cassette exons (Figure 5.B); use of alternative 5' splice sites or alternative 3' splice sites (Figure 5.C and 5.D); retention of alternative introns (Figure 5.E) and alternative splicing in conjunction with the use of alternative promoters (Figure 5.F) or poly(A) sites alternatives (Figure 5.G) (Matlin *et al.,* 2005); (Kornblihtt *et al.*, 2013).

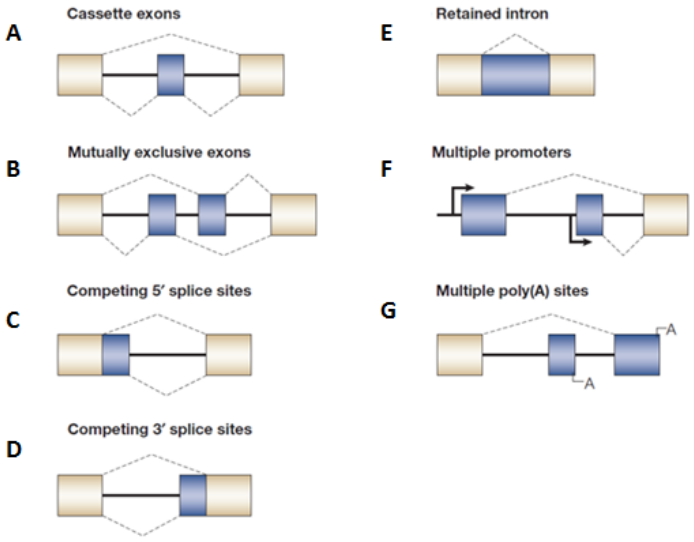


Figure 5. Different modes of alternative splicing. Adapted from (Matlin *et al.,* 2005).

1.6. Splicing and disease

Since splicing is a major regulatory step in gene expression, any alteration of the splicing pattern of one or multiple transcripts can disrupt the production or function(s) of the encoded protein(s) leading to a wide range of phenotypic consequences.

Actually, splicing mutations account for approximately 10% of roughly 80,000 mutations associated to inherited diseases contained in the human gene mutation database (2014). In 2003, Faustino and Cooper have classified this type of disease causing alterations into four

categories: *cis*-acting mutations, those that disrupt the use of constitutive splice sites; *cis*-acting mutations, when they disrupt the use of alternative splice sites; *trans*-acting mutations, if compromising the basal splicing machinery and *trans*-acting mutations, when disturbing splicing regulation.

Cis-acting mutations that affect the use of constitutive splice sites most of the times result in loss of gene expression due to nonsense-mediated decay (NMD) or expression of proteins containing internal deletions, a shift in the reading frame, or C-terminal truncations. In fact, the most commonly detected splicing mutations are located at conserved donor (5') and acceptor (3') splice sites while mutations in the branch site or in the polypyrimidine tract are rare (Lewandowska, 2013). Furthermore, alteration in ESE, ESS, ISE, or ISS have also been recognized as a cause of genetic disease through splicing alteration (Baralle and Baralle, 2005). This type of mutations frequently result in exon skipping, use of nearby cryptic 5´or 3´splice site, or even in retention of the mutated intron. Alternative splice sites can also be mutated and, in this case, a mutation that inactivates one of the alternatively used splice sites will force expression of the alternative splicing patterns.

Trans-acting splicing mutations, in its turn, can affect the function of the basal splicing machinery or factors that regulate alternative splicing. Mutations that affect the basal splicing machinery have the potential to affect splicing of all pre-mRNAs, whereas mutations that affect a regulator of alternative splicing will affect only the subset of pre-mRNAs that are targets of the regulator (Faustino and Cooper, 2003).

The more that is known about RNA splicing, the more is recognized its relevance in terms of health and disease. In figure 6, are illustrated examples of different splicing events and their consequences at protein level.

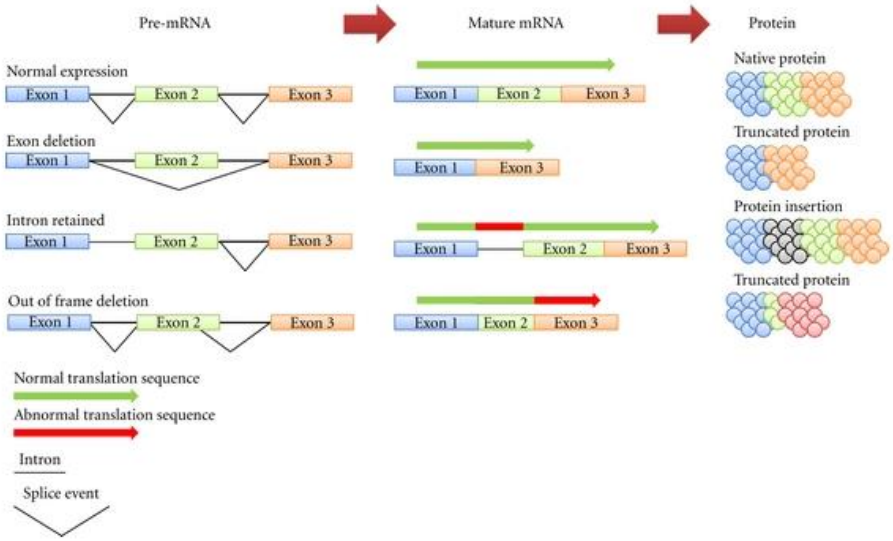


Figure 6. Examples of different splicing events and their consequences at protein level due to changes in pre-mRNA expression that can cause disease. In the process of splicing can happen the removal of exons or retention of introns resulting in the formation of different mature mRNA transcripts for the same gene. Different mature mRNA transcripts encode for different proteins (Faiz and Burgess, 2012).

1.7. Use of minigenes to study splicing mutations

In recent years, the development of high throughput technologies allowed the detection of many changes in intronic sequences and intron/exon boundaries.

The assessment of whether a mutation affects exon recognition and results in a genetic disorder is, however, not always simple. Direct analysis of the processed transcripts from the patients is the most reliable method to confirm that a specific alteration affects splicing and should therefore be the method of choice to investigate the consequences of a splicing mutation. Unfortunately, in many cases it is not possible to perform such analysis as for example, when is not available, a sample from the patient, or from the appropriate tissue when the mutation affects tissue specific gene expression, suitable to extract RNA (Baralle *et al.,* 2009). To overcome these circumstances, the establishment of a causal effect of a mutation can rely on the use of molecular biology methods such as *in vitro* transcription of the sequence of interest cloned into a plasmid, with and without alterations, or hybrid minigene systems to perform mutation analysis (Lewandowska, 2013).

Minigene constructs (Figure 7) have revealed to be an important tool for the identification and *in vivo* analysis of the *cis*-acting regulatory elements and *trans*-acting factors that establish splicing efficiency and that regulate alternative splicing. Expression of minigene pre-mRNAs by transient transfection provides a rapid assay to perform loss or gain of function analyses of factors that affect splicing regulation and therefore can be used to address a number of different questions with regard to alternative splicing and splicing mutations (Cooper, 2005).

When the aim is to study a putative splicing mutation, the minigene should contain a genomic segment from the gene of interest that includes the site of the alteration, usually the exon and part of the flanking introns. This fragment, which will be cloned into splicing vector, is amplified from genomic DNA (from normal and affected individuals) using oligonucleotides containing restriction enzyme sites at their 5' ends that match restriction sites in the recipient plasmid, being most convenient to select two restriction enzymes that cut efficiently in the same buffer to obtain more efficient and reliable reaction (Gaildrat *et al.,* 2010). Alternatively, the mutant minigene can also be constructed through site directed mutagenesis of the control/normal one.

Once constructed, the minigenes are transfected into the appropriate cultured cells where RNA undergoes splicing. The comparison of the splicing patterns, through RT-PCR using vector specific primers, of the transcripts generated from the wild-type and from the mutant minigenes, affords evidence on the functional effect of the alteration. The main strengths of the minigene approach are the abilities to demonstrate that specific nucleotide changes affect splicing efficiency and to define elements required for responsiveness to cell type and specific splicing regulators (Cooper, 2005) (Perez *et al.,* 2013).

Many studies have already been undertaken in which minigenes were used to understand and assess the impact of splicing alterations. For example, in 2010 the strategy was applied by Fernández-Guerra and collaborators to demonstrate the pathogenic impact of an intronic mutation in *BCKDHA* gene causing a variant form of MSUD (Fernández-Guerra *et al.*, 2010).

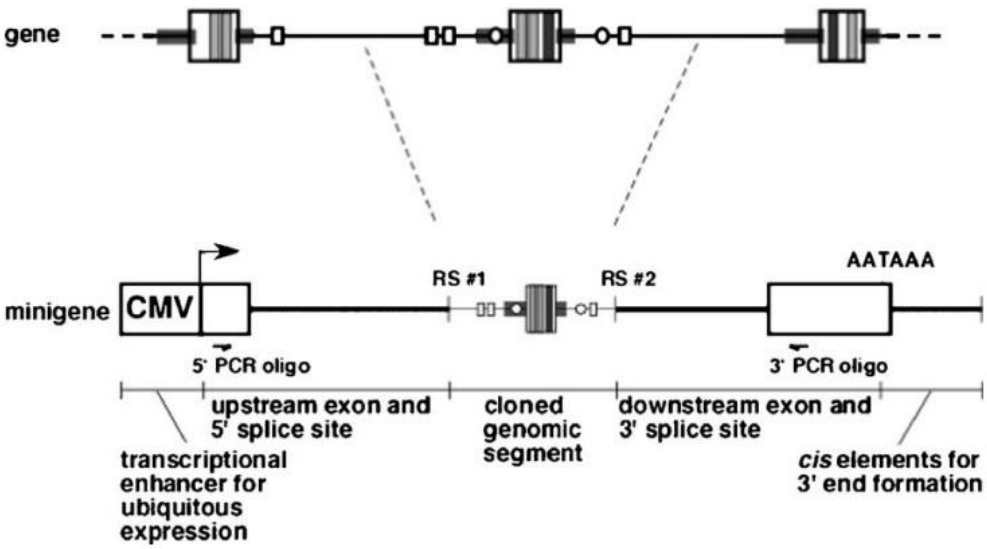


Figure 7. Representative scheme of the construction of a minigene in a plasmid vector. CMV – cytomegalovirus transcriptional enhancer/promoter. RS #1 and RS#2 – restriction sites located on the plasmid. The genomic fragment to be cloned into the minigene is amplified from genomic DNA using oligonucleotides containing restriction enzyme sites at their 5' ends that match restriction sites in the plasmid. These sites depend on the available cloning sites within the minigene (RS#1 and RS#2) (Cooper 2005).

2. Aims

While molecular characterization of MSUD patients is essential to provide carrier testing for at-risk relatives and pre-natal diagnosis for high-risk pregnancies, it might also be important for a proper disease management. On the other hand, understanding the pathobiology of mutations is fundamental to strengthen the relationship between molecular defect and disease.

Acknowledging that, the main objectives of this work were:

1) To perform the molecular characterization of two patients recently diagnosed with MSUD by amplification and sequencing the entire coding and flanking intronic regions of *BCKDHA*, *BCKDHB* and *DBT* genes;

2) To evaluate the putative pathogenic effect of newly identified missense mutation(s) through the use of available bioinformatics programs that predict the impact in the protein structure induced by amino acid replacements;

3) To assess *in silico* tools to predict the effect caused by a new variation identified in intron 1 of *BCKDHA* gene (c.108+6T>C) on the splicing process of the gene;

4) To determine the functional impact of the substitution c.108+6T>C in *BCKDHA* through the construction of a splicing reporter minigene assay;

5) In a broader perspective, it was also an objective of this work to enlarge the knowledge on the molecular architecture of MSUD and on the pathobiology of mutations that underlie the disease.

3. Material & Methods

3.1 - Molecular characterization of MSUD patients

3.1.1- Samples and DNA extraction

In this study, samples from two MSUD patients whose biochemical diagnosis has been recently established have been analyzed to identify the disease causal mutations. Moreover, a sample from a Brazilian patient, for whom the molecular characterization has already been performed, was also studied.

DNA was extracted from patients' blood spot in FTA card using the Generation Capture Card Kit following the protocol "Purifying DNA from difficult-to-elute dried blood spots" provided by the manufacturer (Qiagen).

3.1.2- Amplification by PCR

In order to identify the MSUD causal mutations in the patients, the entire coding and flaking intronic regions of the three genes most frequently associated to MSUD (*BCKDHA*, *BCKDHB* and *DBT*) were amplified by PCR using specific primers previously described in Quental *et al*., 2008 (Quental *et al.,* 2008). The PCR reaction mixtures consisted in 5 µl of MyTaq™ HS Mix (Bioline); 1 µl of primer forward (10 mM); 1 µl of primer reverse (10 mM); 1 µl of DNA and 2 µl of H_2O to complete a final volume of 10 µl. The amplification programs are described in table 1 whereas table 2 present sequences and annealing temperatures of primers used.

Table 1. PCR program used to amplify the entire coding and flaking intronic regions of the three genes that are most frequently affected in MSUD (*BCKDHA*, *BCKDHB* and *DBT*)

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Denaturation	95	15 '	1
Denaturation	94	30 ''	35
Annealing	Variable*	1' 30''	
Extension	72	1'	
Final Extension	72	10'	1
Hold	4	∞	1

*this information is present in table 2 and depends on the pair of primers used

Table 2. Sequences and annealing temperatures of primers used in amplification of entire coding and flaking intronic regions of the three genes that are most frequently affected in MSUD (*BCKDHA*, *BCKDHB* and *DBT*)**.**

Exons	Primers names	Sequences (5'- 3')	Ta (ºC)
		BCKDHA	
1	1F/1R	CAGAACGGGAAGAAGATGGT / CTTCGGTGTCTTCTCCAAGG	59
2/3	116-374F/116-374R	GCTTCTGATGCAGGTGGTCT / CCTCAGGTAATTCCAGCCCC	61
4	4F/4R	CCCAGCATAACCAATTGTGG / GCTGCTCCTGGAAGAACACT	59
5	5F/5R	TCCTCTTAAAACAAGCCTGAGC / TCACCAACCCAGAATTCCAC	61
6	6F/6R	CTGCTCACCACCCTCTCATC / CACAGGACACAGGACGAGAA	60
7	7F/7R	CACCCCTACCCTCCTTCCT / GTGGCTGTCAGTGCTGTGG	61
8	8F/8R	TGCCTTTATTCCGTTTCCAC / ATCACTGGGGTTATCCCTGA	59
9	1234-1312F/1234-1312R	CACAAAGGCTTGGAGTGGTT / CCCTTAGAGTGGGGCTACCT	59
		BCKDHB	
1	1BF/1BR	CCCCTAAAT TTCCAGTTCCG / AATAAGCTGGGATGCAAGGA	55
2	2-3BF/2-3BR	ACTTTGACGGGTCTCCCTTT / TGTTAACCATTGAGCTCCACA	60
3	4BF/4BR	GGTAACTGTCATCCAGTGGGTAG / CTGCGGGTGGCGTTGGAAATG	61
4	6BF/6BR	TGACATTACTCTCATTTGCCACA / GGGTAGCGGCAATACTTGAA	61
5	7BF/7BR	AGGAGATTGGAAGGGAAGGA / TTCCAAGAATCCTGTTTGTTTTC*	60
6	9BF/9BR	AGCCCTTCTTAGCAGCGAGT / CACCTGAGGACAAACTAAATACCA	61
7	10BF/10BR	TGCACAAGTGTCACCTCAGA / TGAGAGCTTCCAAGCACAAC	61
8	11BF/11BR	TCAGCATTCAACTAGTTTTTGAGG / GCCAAAGGTTTCAGGGAAAT	61
9	13BF/13BR	TTTTCTACTGGGATTACAAACCA / GGAATTGCACAAGCATTGAA	61
10	14-1BF/14-1BR	CACATAATAAAACTGGGATCATG/ CGTTAATGTCAGGGGCACAT	61
		DBT	
1	DBT-1F/R	GCTCGTTTCTTCCCTCCCTA / CATGCATCCCTTCACCTCTC	5*64 + 30*62
2	DBT-2F/R	GAGATAAGCCGGTATGGTTGTT / GCCCGGCTAGAAATACAGAG	5*64 + 30*62
3	DBT-3F/R	ATTTCTGCCTCTGCCTGAGA / GGATTCCCACTATCCATTAA	59
4	DBT-4F/R	TCTGAAAGTAAATGCTGGGCTA / GGGACCCAATGACAAAAAGA	58
5	DBT-6F/R	CTGAAGTTAAGTTTTACCTTGTTAC / AATTTCATTGAGCTATTTCATC	52
6	DBT-7F/R	GATACCTGATGGTTACCACATGCAT / TATGCTGACTGAAGTTGAACTTTC	53
7	DBT-8F/R	GACATTAGAGAACCTTCCATT / CAAGAGCAAAACTCTGTCTC	58
8	DBT-9.1F/R	GCGCTTTGAAATCCTCGTTA / TAATCCCAAGGAAGCAGCCT	62
8	DBT-9.2F/R	CACATGGCCTGAAGGTAACAT / GGGTAACATAATCTGCCATACAGC	5*58 + 30*56
9/10	DBT-10/11F/R	ATGGCAGTGAAGGTTGATCC / CCTACAAAAGGAGGGGAAACC	60
11	DBT-12F/R	TTGAGCTCTGAACAAGTGAAGGTT / CAGCACCATATTAAGAAGTCACAC	63

3.1.3- Evaluation of the success of PCR by polyacrylamide gel electrophoresis

The success of the amplification reactions was tested by submitting the PCR products to polyacrylamide gel (T9C5) electrophoresis. The composition of the gel is described at the table 3. The voltage of 280V was applied during 30 minutes. The gel was visualized by silver staining standard protocol.

Table 3. Composition of a polyacrylamide gel

Component	Volume
Acrylamide solution (Acrylamide/Bisacrylamide 19:1 (40%))	3 ml
Ammonium Persulfate (APS) 2.5%	170 µl
Tetramethylethylenediamine (TEMED)	7 µl

3.1.4- Sanger Sequencing and electropherogram analysis

The amplified products were initially purified with ExoSAP-IT® (USA Products) to remove residues of primers and excess dNTPs not used in the PCR. For the purification reaction, 0.5 µl of the enzyme were mixed with 1.5 µl of amplified product and incubated at 37°C for 15 minutes and another 15 minutes at 85°C. Then, the sequencing reaction was prepared adding to each tube 2 µl of the purified PCR product, 1 µl of oligonucleotide primer 5' or 3' at 10 mM,1 µl of BigDye® Sequencing Standard Kit and 1 µl of sequencing buffer (Applied Biosystems). The sequencing reactions were carried out in a 2720 Thermal Cycler (Applied Biosystems) or in a Thermal Cycler (BioRad) with the conditions described in the table 4.

Table 4. Thermal cycler conditions for the sequencing reaction

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Denaturation	96	4'	1
Denaturation	96	10 ''	35
Annealing	55	5''	
Extension	60	4'	
Final Extension	60	10'	1
Hold	8	∞	1

At the end of the reaction, the sequencing products were purified in SEPHADEX® columns (GE Healthcare) through centrifugation for 4 minutes at 4400 rpm to remove all the oligonucleotide primers, dNTPs and ddNTPs that have not been used in the reaction. Finally, the purified products were ressuspended in 12 µl of the denaturing agent Hi-Di formamide and the capillary electrophoresis was performed on an ABI Prism *3130xl* DNA Sequencer (Perkin-Elmer, Applied Biosystems).

The electropherograms obtained were compared to reference sequences available in the on-line database Ensembl Genome Browser - http://www.ensembl.org/index.html, using Geneious 5.5.8 program.

3.1.5- Bioinformatic analysis

To predict the impact of novel missense alterations discarded to be polymorphisms the program PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/) was used.

In silico analysis of a likely splicing mutation was performed using the following bioinformatics programs:

GenScan (http://genes.mit.edu/GENSCAN.html) (Burge and Karlin 1997)
NNSplice: Splice Site Prediction by Neural Network (http://www.fruitfly.org/)
MaxEnt (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html)
Human Splicing Finder (http://www.umd.be/HSF/)
NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/).

3.1.6- Protein structural analysis

The E1β protein structural model with a newly identified mutation was built with MODELLER (Eswar *et al*., 2006), using the normal human sequence (PDB ID: 1X7Y) as template (Wynn *et al*., 2004).

Generation of protein plots was performed using PyMOL (http://www.pymol.org) (DeLano).

3.2- Functional characterization of the alteration c.108+6T>C within *BCKDHA* gene – Construction of reporter minigenes

The functional studies were planned for an alteration, c.108+6T>C, that was located in intron 1 of *BCKDHA* gene, lacking so any upstream 3' splice site. It was needed, therefore, to produce first a rearranged exon resultant from the fusion of two exonic parts (one of them the *BCKDHA* exon 1) in order to obtain a hybrid exon possessing both 3´and 5´splice sites. For this reason, a hybrid splicing vector (pSPL3), already used in previous studies (Arrabal *et al.,* 2011; Perez *et al.,* 2013), that contains part of exon 2 of *SPR* gene fused to exon 1 of *ALDH7A1* (Figure 8) (Pérez *et al.*, 2013) was employed and modified in the present study. It was kindly provided by Bélen Pérez, Madrid, Spain.

Figure 8. Representation of the pSPL3 splicing vector that contains the hybrid exon formed by part of *SPR* exon 2 fused to *ALDH7A1* exon 1. Adapted from Perez *et al*. 2013 (Perez *et al.,* 2013)

In order to modify the vector, the region encompassing exon 1 of *ALDH7A1* was replaced by the region of interest within *BCKDHA*, through a process that is described in detail in the next topics.

3.2.1- Construction of intermediate minigene vector

3.2.1.1- Amplification of the *BCKDHA* genomic DNA fragment covering the target region

A genomic fragment of *BCKDHA* gene including the region where lies the c.108+6T>C mutation (203 bp of exon 1 and 301 bp of intron 1) was PCR amplified from a control DNA sample using a forward primer that includes in the 5' side the restriction site (underlined) for SalI (ATATATGTCGACGTCAGATGGTGCTGGAGTTTG- named BCKDHA_SalI_F) and a reverse primer that includes the restriction site for EcoRV (TAAGCAGATATCCCTCCCCTGAATACCAGTGG – named BCKDHA_EcoRV_R_new) (Figure 9). Since usually restriction enzymes do not cut efficiently at sites nearby the end of DNA fragments, six extra nucleotides have been included as the first nucleotides in each primer (Cooper, 2005). The amplification reaction was performed with HotStar HiFidelity Polymerase Kit (Qiagen) using 10 µl of HotStar HiFidelity PCR Buffer 5X; 4 µl of forward and reverse primers (10 mM); 1 µl of HotStar HiFidelity DNA Polymerase; 2.5

µl of DNA and 28.5 µl of H_2O to complete a final volume of 50 µl. The PCR conditions used are described in table 5.

Table 5. PCR conditions used to amplify a genomic fragment of *BCKDHA* gene

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Activation	95	5'	1
Denaturation	94	15''	35
Annealing	64	1'	
Extension	72	1'	
Final Extension	72	10'	1
Hold	4	∞	1

The success of this reaction was tested through polyacrylamide gel electrophoresis as described in point 3.1.3, and, to confirm the integrity of the amplified fragment, the product was sequenced with the same primers used in the PCR and the conditions described in 3.1.4.

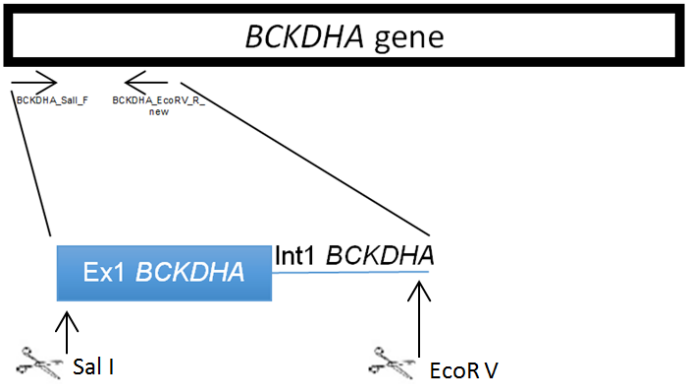


Figure 9. Schematic representation of the fragment amplified from *BCKDHA* gene including the region of the c.108+6T>C mutation (203 bp of exon 1 and 301 bp of intron 1) using primers BCKDHA_SalI_F and BCKDHA_EcoRV_R_new.

3.2.1.2- Amplification of the region of interest of pSPL3 vector

The region of the pSPL3 vector that includes *SPR* exon 2 fused to exon 1 of *ALDH7A1* was PCR amplified with the HotStar HiFidelity Polymerase Kit, using the same reaction conditions described in 3.2.1.1 but with primers pSPL3_E2E1_F (GGATCACCAGAATTCTGGAGC) and pSPL3_E2E1_R

(TCCACACAGGTACGGGATCA) (Figure 10). The conditions of this PCR are described at table 6.

Table 6. PCR conditions used to amplify the region of pSPL3 vector that includes *SPR* exon 2 fused to exon 1 of *ALDH7A1*

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Activation	95	5'	1
Denaturation	94	15''	35
Annealing	61	1'	
Extension	72	1'	
Final Extension	72	10'	1
Hold	4	∞	1

The success of this reaction was also tested in polyacrylamide gel electrophoresis and the amplified product was sequenced with the primers used in the amplification.

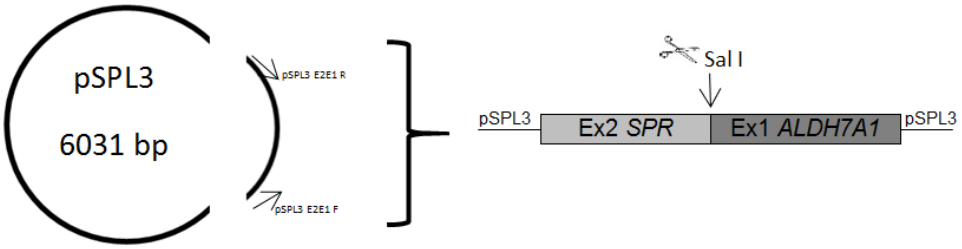


Figure 10. Representation of the fragment amplified from pSPL3 vector that includes part of exon 2 of sepiapterin reductase (*SPR*) gene fused to exon 1 of *ALDH7A1* gene (Pérez *et al.*, 2013) with primers pSPL3_E2E1_F and pSPL3_E2E1_R.

3.2.1.3- Enzymatic restriction of DNA fragments from *BCKDHA* gene and pSPL3 vector with SalI

After confirming the integrity of the pSPL3 and *BCKDHA* fragments by Sanger sequencing, 21 µl of each amplicon was digested (in independent reactions) with restriction endonuclease SalI (Roche) using 1.5 µl of enzyme and 2.5 µl of H Buffer (final volume of 25 µl). The restriction sites for Sal I of the *BCKDHA* and pSPL3 fragments are schematically represented in figures 9 and 10, respectively. The reactions occurred at

37°C, the ideal temperature for the enzyme activity, for 2 hours and 30 minutes, and were performed in duplicate for both fragments.

The products of these restrictions were separated in 2.5% agarose (ultrapure grade) gel electrophoresis and the fragment of pSPL3 containing *SPR* exon 2 and the fragment from *BCKDHA* were gel purified with Wizard® SV Gel and PCR Clean-Up System (Promega) following the manufacturer's instructions. The correspondent bands from the duplicate reactions were eluted together in order to obtain DNA with concentration enough to be used in later procedures.

The success of the enzymatic restriction reactions was confirmed by Sanger sequencing of the eluted DNAs using the same primers as in the PCR amplifications for *BCKDHA* fragment and with primer pSPL3_E2E1_F in the case of pSPL3; the quality of the DNAs was evaluated with NanoDrop ND-1000 Spectrophotometer.

3.2.1.4- Ligation of *SPR* exon 2 to *BCKDHA* fragment

The ligation of *SPR* exon 2 to the fragment from *BCKDHA* gene, which includes exon 1 and part of intron 1, was performed with Rapid DNA Dephos & Ligation Kit (Roche) using 4 µl of each DNA and 2 µl from DNA Dilution Buffer 5X. To this mixture it was added 10 µl of T4 DNA Ligation Buffer 1X and 1µl of T4 DNA Ligase. The ligation reaction was incubated for 60 minutes at room temperature and overnight at 4°C.

The product of the ligation was amplified by PCR with HotStar HiFidelity Polymerase Kit using as primer forward primer pSPL3_E2E1_F and as primer reverse primer BCKDHA_EcoRV_R_new. PCR conditions used are described in table 7.

Then, the ligation product was submitted to 1% agarose gel electrophoresis and the fragment with the expected size was eluted from the gel with Wizard® SV Gel and PCR Clean-Up System (Promega).

The integrity of the ligation fragment amplified was evaluated by Sanger sequencing with primers pSPL3_E2E1_F and BCKDHA_EcoRV_R_new. The quality of the DNA was accessed with NanoDrop ND-1000 Spectrophotometer (Thermo scientific).

Table 7. PCR conditions used in the amplification of the ligation product of *SPR* exon 2 and *BCKDHA* fragment

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Activation	95	5'	1
Denaturation	94	15''	35
Annealing	64	1'	
Extension	72	1'	
Final Extension	72	20'	1
Hold	4	∞	1

3.2.1.5- Cloning of PCR product in TOPO TA cloning vector (pCR®2.1-TOPO®)

The fresh PCR product eluted from the gel (3.2.1.4) was cloned in TOPO TA vector (Invitrogen) (Figure 11). The reaction mixture containing 4 µl of amplified product; 1 µl of salt solution and 1 µl of TOPO® vector in a final volume of 6 µl, was incubated during 30 minutes at room temperature.

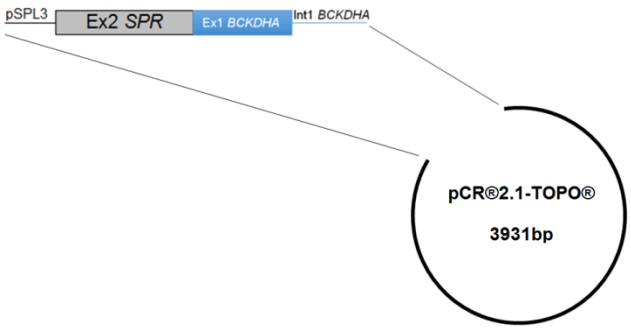


Figure 11. Cloning of the hybrid vector (SPR-BCKDHA) in TOPO TA cloning vector.

3.2.1.6- Transformation of One Shot® TOP 10 Chemically Competent Cells

The cloning reaction was used to transform TOP10 Chemically competent cells (One Shot® TOP10 Chemically Competent *E. coli*, Invitrogen). The transformation was performed through the addition of 2 µl of the cloning product to a vial with 50 µl of TOP 10 *E. coli* cells. This mixture was incubated on ice for 30 minutes and then heat-shocked for 30 seconds at 42ºC, and immediately transferred to ice during 2 minutes. Finally, 250 µl of pre-heated S.O.C. medium were added before incubating the mixture at 37ºC with horizontal agitation at 200 rpm for 1 hour.

Different volumes of transformation product (50 µl and 100 µl) were plated on pre-heated LB-Agar plates with ampicillin (50 µg/ml), and incubated at 37ºC during 16-18 hours.

Since the TOPO vector confers antibiotic resistance, expectedly only bacteria that incorporate it should grow and form colonies.

After the incubation period, several individual colonies from both plates were picked with a pipet tip and dipped into 20 µl of H_2O. Colony PCR was then performed using the picked samples to identify the colonies that had incorporated the TOPO vector. This amplification reaction was performed with QIAGEN Multiplex PCR Kit (conditions referred on table 8) using TOPO specific primers M13_F (GTAAAACGACGGCCAG) and M13_ R (CAGGAAACAGCTATGAC). The amplicons were analyzed in polyacrylamide gel electrophoresis to identify putative positive colonies which are those showing a fragment of 906 bp. All amplified products from the suspected positive colonies have been sequenced (with primers M13F and R) to confirm the presence of the insert and if it was in the correct orientation.

Positive colonies were maintained in culture using the remaining volume of solution not spent in the PCR and the pipet tip used to pick the colony and both were dipped into 5 ml of liquid LB medium with ampicillin (50 µg/ml) and cultured overnight at 37°C with agitation at 200 rpm. For long term storage, a glycerol stock was prepared through the addition of 0.15 ml of sterile glycerol to 0.85 ml of the culture, and stored at - 80°C.

Table 8. Colony PCR conditions used to identify positive colonies

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Activation	95	15'	1
Denaturation	94	30''	35
Annealing	54	1'30''	
Extension	72	1'	
Final Extension	72	1'	1
Hold	4	∞	1

3.2.1.7- Miniprep of bacterial cultures

After confirming which colonies had the intermediate minigene vector, DNA minipreps were prepared using the ZYMiniprep kit (NZYTech), following the manufacturer's instructions.

Isolated DNAs were quantified in NanoDrop ND-1000 Spectrophotometer (Thermo scientific) and sequenced with primers M13F, M13R and BCKDHA_EcoRV_R_new to confirm that no alterations had been introduced.

3.2.2- Construction of final minigene splicing vectors

3.2.2.1- Digestion of pSPL3 splicing vector and intermediate minigene vector with BamHI and EcoRV

To construct the final expression vector, both intermediate TOPO and pSPL3 vectors were digested with BamHI and EcoRV restriction enzymes (both from Roche) (Figure 12). As recommended in the protocol, 3 µg of each DNA were mixed with 0.6 µl of BamH I; 0.6 µl of EcoR V; 3 µl of H Buffer and 15.8 µl of H_2O to perform a final volume of 30 µl. The digestion reactions occurred at 37°C for 2 hours and 30 minutes.

The restriction products were separated in 1% agarose gel electrophoresis and after conventional DNA staining, the band from pSPL3 thought to correspond to the vector without the “original” insert and the band assumed to be the fragment that includes the hybrid exon *SPR-BCKDHA* from TOPO, were eluted from the gel with Wizard® SV Gel and PCR Clean-Up System kit (Promega). The quality of the DNAs was evaluated with NanoDrop ND-1000 Spectrophotometer (Thermo scientific).

To confirm the success of the digestions, Sanger sequencing was performed using for pSPL3 DNA the vector specific primers SD6 (TCTGAGTCACCTGGACAACC) and pSPL3_R (CACATGGCTTTAGGCTTTGA), while for the *SPR-BCKDHA* insert the primers were BCKDHA_EcoRV_R_new, previously used, and BCKDHA_int_seq (GCTCGGGGACTGGCTAGAT).

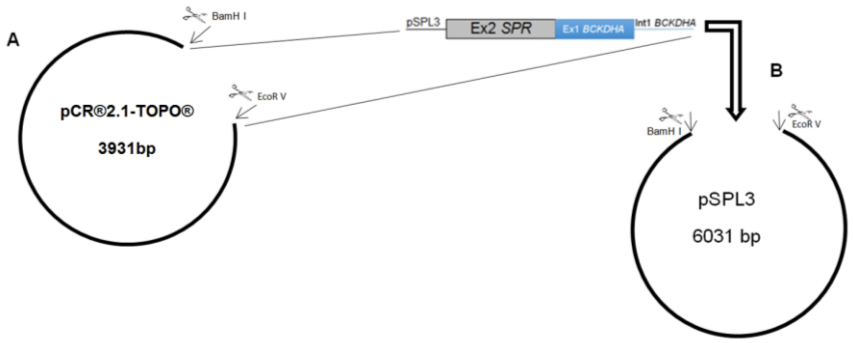


Figure 12. Schematic representation of construction of final minigene vector (insert of interest + pSPL3 vector). **A.** Enzymatic digestion of TOPO vector with restriction enzymes BamH I and EcoR V to extract the insert of interest (part of of sepiapterin reductase (*SPR*) gene fused to part of exon 1 and part of intron 1 of *BCKDHA* gene). **B.** Ligation of insert extracted from TOPO vector to pSPL3 vector after pSPL3 vector has been also digested with restriction enzymes BamH I and EcoR V (mentioned in 3.2.2.2)

3.2.2.2- Ligation of *SPR-BCKDHA* fragment into pSPL3 vector

To construct the final minigene vector, the hybrid exon formed by *SPR* exon 2 and *BCKDHA* exon-intron 1 needed to be cloned in the pSPL3 vector. To that end, 150 ng of the insert were mixed with 50 ng of pSPL3 “empty” vector; 1 µl of T4 DNA Ligase (5 U/µl); 2 µl of reaction buffer 10X; and H_2O to obtain a total volume of 20 µl. Finally, the mixture was incubated at 18°C for 16 hours and at 65°C for 10 minutes.

3.2.2.3- Transformation of TOP10 Chemically competent Cells with the minigene vector

The ligation product (2.5 µl) was used to transform One Shot® TOP10 Chemically Competent *E. coli* cells (25 µl), following the protocol previously described in 3.2.1.6.

Colony PCR at 57ºC of annealing temperature was performed to select positive colonies, using the pSPL3 specific primer SD6 (TCTGAGTCACCTGGACAACC) and primer BCKDHA_EcoRV_R (ATATATGATATCAGCCCTTCCTTCATATCCTTC), that anneals to the *BCKDHA* part of the insert .

Positive colonies, for which, as expected, a band of 979 bp was observed in polyacrylamide gel, were sequenced with primers used in the PCR, to confirm integrity of the fragment inserted. Bacterial culture, DNA miniprep and glycerol stock have been performed as described in 3.2.1.6 and 3.2.1.7.

3.2.3- Mutagenesis

The mutation c.108+6T>C in intron 1 of *BCKDHA* gene was introduced in the wild-type minigene vector by site directed mutagenesis using the QuikChange II Site-Directed Mutagenesis Kit (Stratagene). The QuikChange Primer Design from Agilent Technlogies was used to design the primers sense and antisense with the mutation, which were: BCKDHA_MUT_108_6_C_Fw (GCTAGATCTGTGAGCACCTGGGCCCCAGG) and BCKDHA_MUT_108_6_C_Rv (CCTGGGGCCCAGGTGCTCACAGATCTAGC).

Mutagenesis reaction was performed according to the manufacturer´s protocol that briefly involves the mixture of 2.5 µl of 10X reaction buffer; 0.5 µl of dNTPs; 1.0 µl of each primer at 65 ng/µl; 0.75 µl of DMSO; 0.5 µl of *Pfu*Ultra High-Fidelity; 50 ng of template DNA and H_2O to complete a final volume of 25 µl. The mixture was subjected to 1 minute at 95ºC, followed by 12 cycles of 50 seconds at 95ºC, 50 seconds at 58ºC, and 7 minutes at 68ºC. After the PCR, 0.75 µl of Dpn I was added and the mixture was incubated during 1 hour and 30 minutes at 37ºC.

At the end of the procedure, 2.5 µl of the digested product was used to transform 25 µl of XL-1 Blue supercompetent cells, as recommended by the manufacturer.

Positive colonies were identified by colony PCR and Sanger sequencing (as previously described for WT minigene colonies in 3.2.2.3) to confirm the presence of the mutation c.108+6T>C and the absence of other alterations inadvertently introduced. Bacterial culture, DNA miniprep and evaluation of the quality of the DNA obtained were performed as previously described.

3.3. Cell culture and transfection

3.3.1- Culture and cell counting

In this work, two adherent cell lines- HeLa (human cervix epithelioid carcinoma) and AGS (human gastric carcinoma) - were selected to be transfected with the minigene vectors.

The entire procedure was performed in a laminar flow hood and all the necessary non-sterile material was disinfected with 70% ethanol to prevent contamination of the cells.

Cryopreserved HeLa and AGS cells were properly resuspended to be then maintained in culture with *Dulbecco's Modified Eagle Medium* (DMEM) and Roswell Park Memorial Institute medium (RPMI), respectively, supplemented with 10% Fetal Bovine Serum (FBS) and 1% Penicillin Streptomycin (Pen Strep). The cells were maintained in an incubator at 37°C in a humidified atmosphere (95%) and of 5% CO_2, with medium regularly renewed. Confluence and viability were always observed after examination with an inverted microscope.

When the cell culture was confluent and showed uniform distribution in flasks, the culture medium was removed, cells were washed with phosphate buffered saline (PBS) and trypsin-EDTA was added to dissociate cells from culture flask. After 5-10 minutes at 37°C, when the cells were already detached from the wall, they were sub-cultured by transferring the appropriate volume to a new flask with fresh medium.

In order to transfect cells, a total of $4x10^5$ cells were plated in each well of a 6-weel plate. To determine the appropriate volume of cell suspension that should be used in each well, cells were counted with a Neubauer chamber (Figure 13).

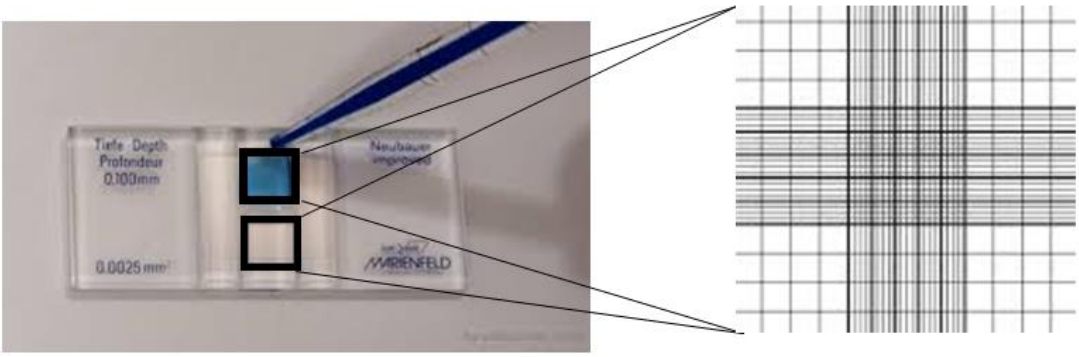


Figura 13. Example of a Neubauer chamber to count cells.

To do that, a solution with 10 µl of cells suspension and 10 µl of 0.1% trypan blue, which also allows checking for cell viability, was prepared. Ten microliters of this mixture was placed in each side of the chamber (Figure 13) and the cells located in the four external quadrants were counted in an inverted microscope. Each square has an area of 1 mm^2, a depth of 0.1 mm, and so, a volume of 0.1 mm^3.

The calculation of the cell concentration was determined by the following expression:

$$N = \frac{\sum C}{4} \times D \times 10^4 \text{ cells/ml},$$

where N is the number of cells per ml; C is the number of cells counted in each quadrant and D is the dilution factor which in this case was 2 (10 µl of cells suspension diluted in 10 µl of Trypan blue). The value obtained was multiplied by 1000 to determine cell count in ml.

The 6-well plate was incubated at *5*% CO_2, 95% humidity and 37°*C.* After 24h, the cells were at a confluence of about 70-80% and therefore suitable to be transfected.

3.3.2- Transfection

Three different plasmid DNAs were used to transfect both cell lines (HeLa and AGS), which were the empty pSPL3 vector, the WT minigene (c.108+6T) and the mutant minigene (c.108+6C) vector, as shown in figure 14.

Cells were transfected with each DNA using JetPEI® transfection reagent following manufacturer's instructions (POLYPLUS TRANSFECTION). Briefly, 2 µg of each vector DNA and 8 µl of JetPEI were independently mixed with 100 µl of 150 mM NaCl; the mixtures of JetPEI were then added to each DNA diluted one. Each of these solutions was incubated at room temperature for 30 minutes after which 200 µl of each was delivered to the respective well containing the cells growing in 2 ml of supplemented DMEM or RPMI medium (Figure 14).

The plate was incubated at 5% CO_2, 95% humidity and *37*°C for approximately 24 hours until cell recover for total RNA isolation.

Three independent transfection experiments were performed for each plasmid tested to assess for reproducibility of the results.

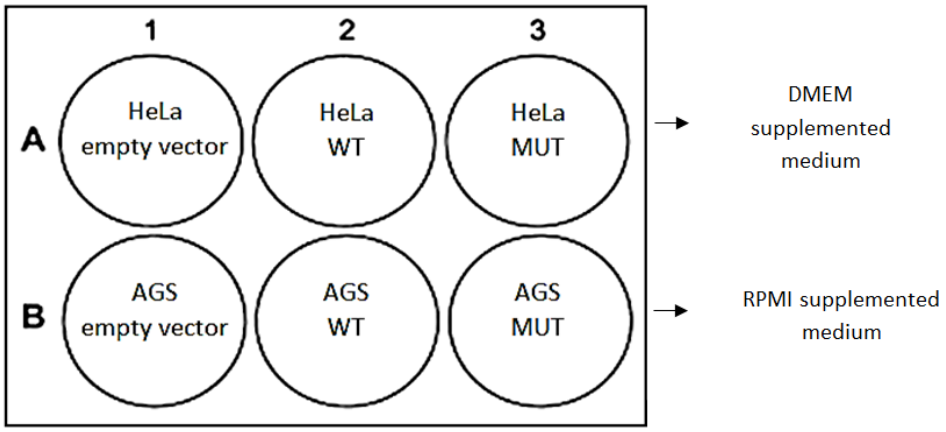


Figure 14. Representation of transfection in the 6-well plate. In A1 it is present HeLa cells transfected with empty pSPL3 vector; in A2 it is present HeLa cells transfected with wild-type (WT) vector and in A3 it is present HeLa cells transfected with vector with the mutation c.108+6T>C in *BCKDHA*. In line B is the same but for AGS cells.

3.3.3- RNA isolation

Approximately twenty-four hours after transfection, the medium was removed, and the cells were washed with PBS and harvested by trypsinization as described above in 3.3.1. Then, 1 ml of DMEM supplemented medium was added to wells containing HeLa cells and 1 ml of RPMI supplemented medium to the ones containing AGS cells. The entire volume was collected into 1.5 ml tubes that were centrifuged at 1500 rpm for 5 minutes to remove the supernatant. Finally, the cells were washed with 1 ml of PBS 1X and centrifuged again at 1500 rpm for 5 minutes.

RNA was extracted from the cell pellets with the High Pure RNA Isolation Kit (Roche), following manufacturer's recommendations.

The quality of all RNA samples was assessed using NanoDrop ND-1000 Spectrophotometer (Thermo scientific).

3.3.4- cDNA synthesis by reverse transcription

The reverse transcription reactions to synthesize cDNAs were performed with the First Strand cDNA Synthesis Kit (Fermentas). All the components of these reactions are shown in table 9.

Table 9. Components for the reverse transcription reaction

RNA	**1-1.5 µg (depending on the concentration of RNA extracted)**
Oligo (dT) primers	1 µl
H_2O	To perform 11 µl
Total volume	11 µl
5X Reaction Buffer	4 µl
RiboLock Rnase Inhibitor (20 u/µl)	1 µl
10 mM dNTP Mix	2 µl
RevertAid-M-MuLV Reverse Transcriptase (200 u/µl)	2 µl
Total volume	20 µl

The mixtures were subsequently incubated at 37°C for 60 minutes followed by 5 minutes at 70°C. Control reactions were also prepared, namely a reverse transcriptase negative control to assess for genomic DNA contamination of the RNA samples, which contained every reagent used except the RT enzyme, and a template negative control that contained every reagent except RNA template, allowing to control for reagents contamination.

3.3.5- Transcript analysis

Subsequently to cDNA synthesis, the samples were directly used in PCR reactions using pSPL3 vector specific primers (SD6 and SA2). For the amplification reaction, a final volume of 50 µl was used including 25 µl of 2x QIAGEN Multiplex PCR master mix (Qiagen); 5 µl of each primer (2 µM); 5 µl of cDNA template and 10 µl of H_2O. The PCR conditions are described in table 10.

Table 10. PCR conditions reaction for amplification of cDNA samples

Cycling step	Temperature (°C)	Time	Number of cycles
Initial Denaturation	95	15 '	1
Denaturation	94	30 "	35
Annealing	60	1' 30"	
Extension	72	1'	
Final Extension	72	10'	1
Hold	4	∞	1

The products of the amplification were run in 2% agarose gel electrophoresis, in order to compare the transcripts resultant from the WT and the mutant minigenes. The stained bands in the gel were eluted using the Wizard® SV Gel and PCR Clean-Up System (Promega) and sequenced with primers SA2, SD6 and BCKDHA_int_seq.

4. Results and discussion

4.1. Molecular characterization of MSUD patients

In this study the molecular characterization of two MSUD patients has been successfully achieved. In both patients the entire coding and flaking intronic regions of *BCKDHA*, *BCKDHB* and *DBT* genes were screened to search for the causing mutations of the disease. A third patient (patient 3) was also studied to functionally confirm the pathogenic effect of a substitution previously identified.

Although the BCKD complex, whose enzymatic activity is disrupted in MSUD patients, is composed for more subunits other than the ones encoded by the above mentioned genes, namely the regulatory kinase and phosphatase, experience to date shows that the majority of the disease causal mutations affects one of the three catalytic components specific of BCKD, as it was also corroborated by the results here obtained.

4.1.1. Patient 1

Patient 1, of Angolan origin, presented the usual symptoms of MSUD in the first days of life and the clinical diagnosis was then confirmed. No detailed description of the biochemical characterization and clinical presentation of this patient was available, being, however, known that at the time of the diagnosis he presented with a severe MSUD phenotype.

Sequence analysis revealed the presence of an 11-bp deletion in homozygosity in the *BCKDHB* gene (c.93_103del11) (Figure 15), leading to an alteration of the reading frame starting at residue 32, in which an alanine is replaced by a phenylalanine, and to the insertion of a stop codon 48 residues downstream (p.Ala32Phefs*48). This deletion occurs in exon 1 of the gene, in a region encoding the mitochondrial targeting leader peptide.

The patient here studied manifested the most severe form of the disease which is in accordance with the serious effect caused by the mutation he carries. The alteration leads to the complete loss of function of the E1β protein and to the consequent degradation of E1α because both subunits cannot assemble to form a stable E1 tetramer. It has been suggested that deletions occurring in repetitive genomic regions might often arise due to slipped mispairing or unequal chromosome crossing-over event (Nobukuni *et al.,* 1991). Since the deletion here identified occurs in a repetitive region and in

addition a duplication of the same 11bp sequence was previously identified as a cause of MSUD (Rodriguez-Pombo *et al.,* 2006), it appears that region in the *BCKDHB* gene is indeed prone to deletion events, which probably result from one of the referred mechanisms.

Of note that this alteration was firstly reported in 1991 by Nobukuni and collaborators in Japanese patients and, since then, it was also identified in MSUD individuals from other origins such as Italian, Spanish, Lebanese and now, in this study, Angolan, (Nobukuni *et al.,* 1991) (Parrella *et al.,* 1994) (Rodriguez-Pombo *et al.,* 2006) (Tabbouche *et al.*, 2014). Given the worldwide distribution of c.93_103del11 it would be interesting to perform haplotypic analyses to clarify whether it arose recurrently in different regions, which would pinpoint, if confirmed, the presence of a mutational hot-spot in exon 1 of *BCKDHB* gene.

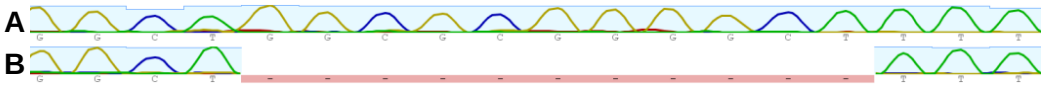


Figure 15. Representation of the 11 bp deletion in *BCKDHB* gene that causes MSUD in patient 1 in comparison with a reference sequence. **A.** Reference sequence. **B.** Sequence of *BCKDHB* gene from patient 1, in which the deletion c.93_103del can be observed.

4.1.2. Patient 2

Patient 2, of Portuguese origin, was identified as a suspect MSUD case in the newborn screening program. The clinical diagnosis made at the initial days of life was further biochemically confirmed (detailed data not available).

At the molecular level, the patient was found to be a compound heterozygous for two mutations also located in *BCKDHB* gene. One of them is a nonsense substitution in exon 7 of the gene, c.799C>T; p.Gln267Ter (Figure 16). This alteration had already been described in patients from different origins (Bentley *et al.,* 2008) but it seems to be particularly common in the Iberian Peninsula. In the study of Rodriguez-Pombo *et al*. (Rodriguez-Pombo *et al.,* 2006), in which 33 Spanish patients were studied, the mutation p.Gln267Ter, occurred in 3 of the 30 *BCKDHB* mutated chromosomes, having been also reported as one of the most frequent mutations in *BCKDHB* gene in Portuguese patients (Quental *et al.,* 2008).

All of the patients carrying the alteration presented with the classic MSUD form, a phenotype which is understandable given the predicted effect of the mutation in impairing complex formation due to the induced E1α and E1β instability that leads to consequent degradation of the committed complex (A *et al.,* 2000).

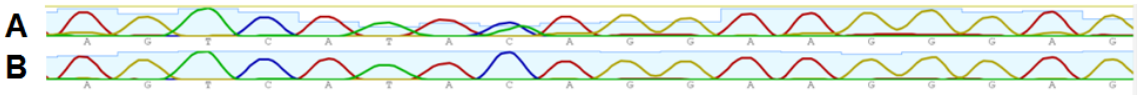


Figure 16. Representation of mutation c.799C>T in *BCKDHB* gene in comparison to a reference DNA. **A.** Sequence of from patient 2, were the mutation c.799C>T is present. **B.** Reference sequence.

Besides p.Gln267Ter, one novel missense alteration also in the *BCKDHB* gene was identified in this patient: c.359T>C, which results in the amino acid change p.Phe120Ser (Figure 17).

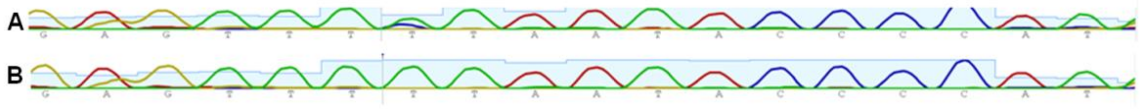


Figure 17. Representation of the mutation c.359T>C in *BCKDHB* gene in comparison to a reference sequence. **A.** Sequence from patient 2, were the mutation c.359T>C in present. **B.** Reference sequence.

Since this alteration was still neither described nor listed as a polymorphism in available databases like Ensembl Genome Browser (http://www.ensembl.org/index.html), several analyses were conducted to assess its putative pathogenic effect.

Firstly, PolyPhen-2 was used, a software tool that predicts the possible impact of an amino acid substitution on the structure and function of a human protein using physical/structural and comparative/evolutionary considerations (Adzhubei *et al.,* 2010). The predictions of the program discriminate substitutions that are probably damaging; possibly damaging or benign according to scores evaluated as 0 (most probably benign) to 1 (most probably damaging).

According to PolyPhen-2, the alteration p.Phe120Ser was functionally damaging for the protein with a score of 1, which corresponds to the highest probability of a substitution be damaging (Figure 18).

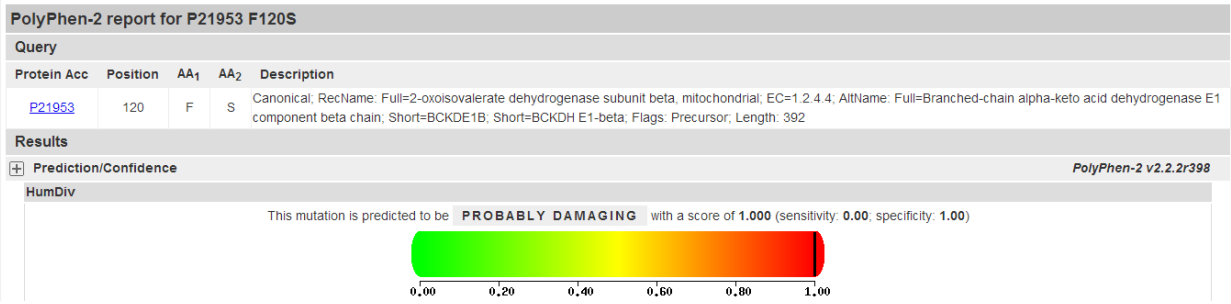


Figure 18. Prediction made by PolyPhen-2 program. It is predicted that the mutation in the Phe120-β residue, p.Phe120Ser, is functionally damaging with a score of 1.

Next, the mutation was analysed with Site Directed Mutator (SDM), which is a method developed to predict changes in the stability of proteins caused by mutations (Topham, Srinivasan *et al*., 1997). SDM calculates a stability score that is analogous to the free energy difference between a wild-type and mutant protein. The stability score is useful for predicting whether a mutation will impact protein structure, having so a role in disease.

SDM uses the local structural environment of the wild-type and mutant residues to calculate the stability score. This score (Pseudo DELTA DELTA G) can be negative, pointing to a destabilizing mutation, or positive, indicating that the mutation is stabilizing. The result obtained using SDM to test the effect of the replacement of phenylalanine by a serine at position 120 of E1β protein (Figure 19), showed a Pseudo DELTA DELTA G of -2.74 for the referred mutation. This score means that this substitution is very probably highly destabilizing, causing protein malfunction and disease.

SDM RESULTS:

Wild-type residue: F
Residue position in wild-type pdb file: 70
Mutant residue: S

Local structural environment of wild-type residue
Secondary structure = extended strand
Solvent accessibility = 26.6% (partially accessible)
Sidechain hydrogen bond satisfaction = NO_HBONDS

Local structural environment of mutant residue
Secondary structure = extended strand
Solvent accessibility = 47.5% (partially accessible)
Sidechain hydrogen bond satisfaction = NO_HBONDS

Pseudo DELTA DELTA G = -2.74

This mutation is predicted to be highly destabilizing and cause protein malfunction and disease.

Figure 19. Predicted effect of the mutation p.Phe120Ser in E1β protein.

To gain insights into the specific structural alterations induced by the referred amino acid replacement, the structure of E1β subunit with the serine residue at position 120 was constructed and compared with the normal one, that is, instead with phenylalanine.

This substitution is located at strand c in the interior of β subunit, which therefore does not seem to affect the interaction with E1α (Figure 20). However, the detailed structural analysis of both E1β proteins (normal and with the mutation) (Figure 21A and B, respectively) revealed that the replacement indeed causes a conformational alteration, because the interactions that residue Arg111 of helix 2 establishes with Lys116, positioned in the loop between helix 2 and strand c, and Val119 from strand c, were no longer observed. The side chains of the mentioned residues are positioned differently as a consequence of the Phe120Ser substitution, preventing these two interactions to be established (Figure 21). It is therefore predictable that the stability of the region involving helix 2, strand c (where the mutation is present) and the loop that exists between these two structures, will be destabilized ultimately affecting BCKD complex activity.

Interestingly, it was before reported the alteration Val119-β, involving the residue immediately upstream of the altered residue here identified (Chuang *et al*., 2004). In the case of Val119-β, the replacement of valine by glycine was predicted to lead to the misfolding of the β subunit. The proximity of the two mutations seems to strength the evidence that the alteration newly identified in this study is the causal mutation responsible for the MSUD phenotype in patient 2.

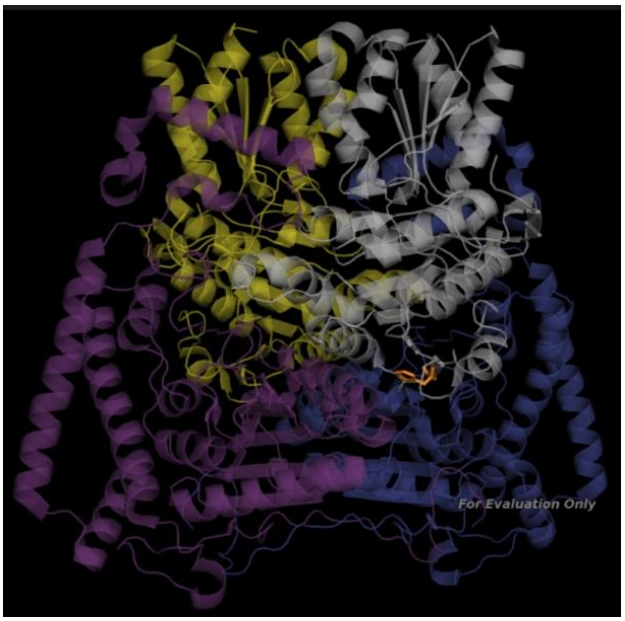


Figure 20. Structural representation of E1 heterotetramer: E1α (purple); E1β gray); E1α′(blue) and E1β′ (yellow). The phenylalanine residue that undergoes the alteration is represented in orange.

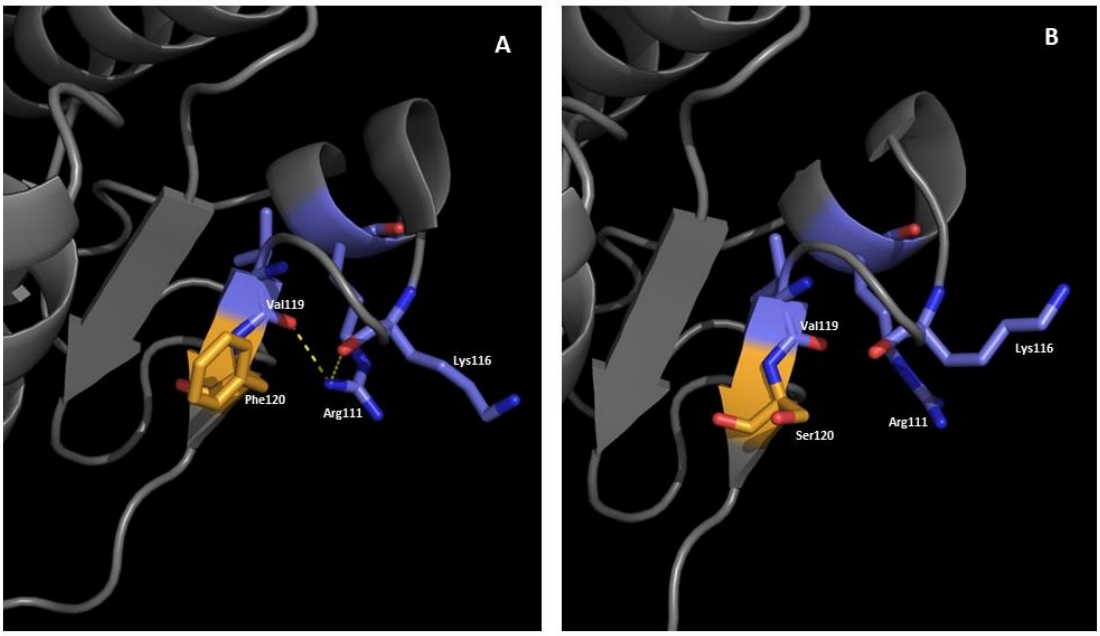


Figure 21. Structural representation of part of E1 β subunit of BCKD complex **A.** Structural representation of normal E1 β protein (120Phe) in which it is possible to observe the interactions that Arg111 located in helix 2, establishes with Lys116, positioned in the loop between helix 2 and strand c, and Val119, situated in strand c. **B.** Structural representation of mutant E1 β protein (120Ser) in which the referred two interactions are no longer established.

4.1.3. Patient 3

The molecular characterization of patient 3, a Brazilian child, had already been performed in our group (data not published). The unique alteration identified in the patient after sequencing *BCKDHA*, *BCKDHB* and *DBT* genes was a mutation in intron 1 of

BCKDHA, involving a substitution of a thymine by a cytosine at position +6 (c.108+6T>C) (Figure 22). Given that this alteration had never been described before, it was quite challenging to ascertain its molecular consequences and pathogenicity, which explains having elected the issue as one of the main objectives of this study. An *in silico* splicing analysis was initially conducted and then a splicing reporter minigene system was implemented. These results for the functional characterization of the alteration c.108+6T>C are described in detail in the next point.

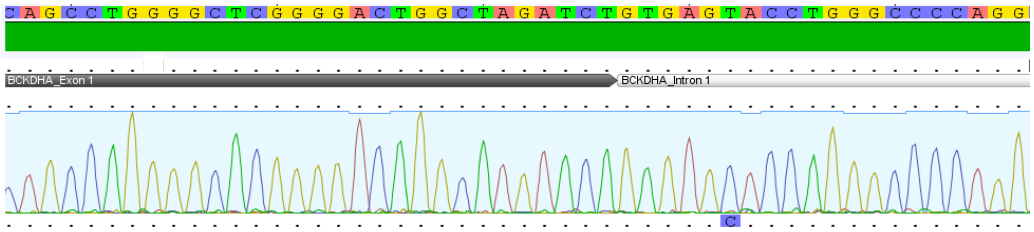


Figure 22. Representation of the mutation c.108+6T>C in intron 1 of *BCKDHA* gene. The substitution in highlighted in blue.

The variations identified in the above mentioned patients are described in table 11.

Table 11. Variations identified in *BCKDHA* and *BCKDHB* genes

Patient	Gene	Exon/Intron	Nucleotide change	Protein prediction	Reference
1	*BCKDHB*	Exon 1	c.93_103del	p.Ala32Phefs*48	(Nobukuni, Mitsubuchi *et al*. 1991)
2	*BCKDHB*	Exon 4	c.359T>C	p.Phe120Ser	This study
2	*BCKDHB*	Exon 7	c.799C>T	p.Gln267Ter	(Nellis *et al.*, 2003)
3	*BCKDHA*	Intron 1	c.108+6T>C	Skipping of exon 1*	This study

DNA mutation numbering is based on cDNA reference sequence (GenBank Accession number NM_ 000709 for *BCKDHA* gene and NM_ 000056 for *BCKDHB* gene) considering nucleotide +1 as the A of the ATG translation initiation codon.
* Results predicted in this study and present in the next topics of Results and Discussion section.

4.2. Functional characterization of the alteration c.108+6T>C within *BCKDHA* gene

The demonstration of the pathogenic impact of an intronic alteration especially when it is not located in the most conserved dinucleotides at each end of the intron, is usually not easy being even more difficult when RNA sample from the carrier individual(s) is not available, as occurs in the case of patient 3 presented in this study.

However, the combination of *in silico* predictions provided by bioinformatics programs with artificial systems using exon-trapping vectors, can afford valuable evidence to elucidate the effect of substitutions in mRNA splicing. Both approaches were here implemented to assess the pathogenicity of c.108+6T>C in *BCKDHA***,** as will be presented in detail in the next topics.

4.2.1. In silico splicing analysis

Five bioinformatics tools were used to evaluate the potential impact of the newly identified mutation c.108+6T>C located in intron 1 of *BCKDHA* gene. Their predictions are hereafter described.

4.2.1.1. GENSCAN (http://genes.mit.edu/GENSCAN.html)

GENSCAN is a bioinformatic program which identifies complete exon/intron structures of genes in genomic DNA. This program is also able to predict multiple genes in a sequence, to deal with partial as well as complete genes and to predict consistent sets of genes occurring on either or both DNA strands (Burge and Karlin, 1997).

In this study, GENSCAN was used giving as input files part of the genomic sequence of *BCKDHA* gene, either the normal sequence or the mutant one, including exon 1 and flanking non-coding regions since c.108+6T>C is located in intron 1. In this way, can be compared the predictions and associated scores on exon 1 given for both sequences.

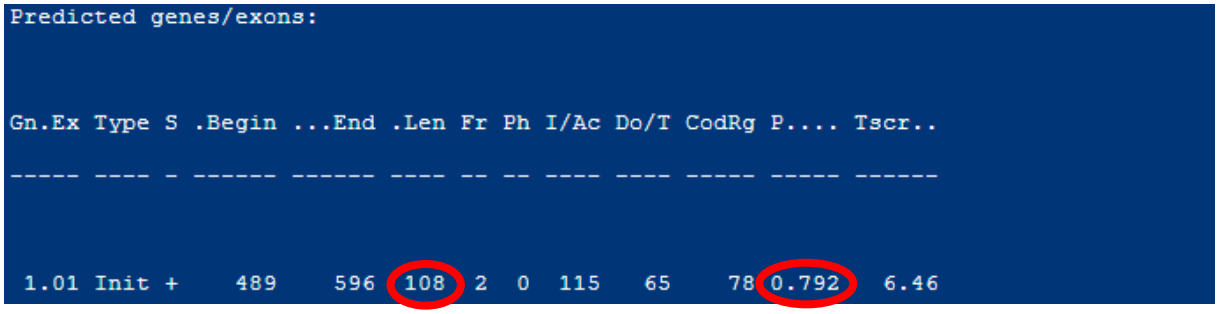


Figure 23. Predicted exon with GENSCAN program, analyzing the normal sequence. The length of predicted exon and the score are surrounded by a red circle.

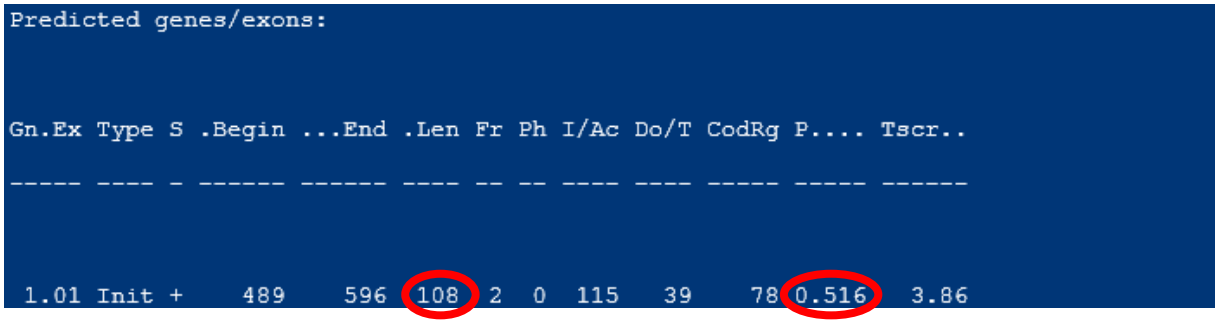


Figure 24. Predicted exon with GENSCAN program, analyzing the sequence with mutation c.108+6T>C. The length of predicted exon and the score are surrounded by a red circle.

Concerning the normal sequence, the program predicted, with the high score of 0.792, the presence of exon 1 in *BCKDHA* encompassing a coding region of 108bp in size (Figure 23), which was correct since the prediction matched the annotated *BCKDHA* sequence. When the sequence with the mutation was used, the program also predicted an exon with the same length, but the yielded probability decreased to 0.516 (Figure 24). This "P" value provides an indication of the likelihood that a predicted exon is correct.

Although these results raised the suggestion that c.108+6T>C might influence the normal splicing pattern, since splicing of exon 1 was not predicted with the same level of certainty in the mutated and normal sequences, they were not clear enough to draw confident conclusions on the effect of this substitution.

4.2.1.2. Splice Site Prediction by Neural Network (NNSplice)

Another bioinformatics resource used was NNSplice (http://www.fruitfly.org/seq_tools/splice.html/), which is a program that leads to splice site predictions for human and *Drosophila melanogaster* by training and testing sets of splice sites that are available in the BDGP (Berkeley Drosophila Genome Drosophila Genome Project) and from the collection of representative and standardized data sets of human and *D. melanogaster* genes. The Splice Site Prediction by Neural Network is a prediction method for donor and acceptor splice sites based on neural networks (Reese *et al.,* 1997).

The prediction that NNSplice produced for the normal sequence of *BCKDHA* was a wild-type 5´ splice site with a score of 0.66. Importantly, giving as input file the mutant

sequence this splice site was no longer identified (cutoff 0.40), indicating that the mutant 5' splice site is probably not used.

4.2.1.3. MaxEntScan

MaxEntScan software uses a method based on the 'Maximum Entropy Principle' to score 5' and 3´splice sites - (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseg.html) (Yeo and Burge, 2004). The higher the score, the higher the probability that the sequence is a true splice site, with ideal scores of 11,81 and 13,59 for 5´and 3´splice sites, respectively (Eng *et al.*, 2004). The utility and accuracy of MaxEntScan was clearly evidenced in the study of Eng, Coutinho *et al.,* 2004, which involved the prediction of splicing mutations in the *ATM* gene (Eng *et al.*, 2004).

```
TCTGTGAGT        MAXENT: 7.00
```

Figure 25. Prediction of maximum entropy with normal sequence

```
TCTGTGAGC        MAXENT: 3.27
```

Figure 26. Prediction of maximum entropy with sequence with mutation c.108+6T>C

Comparing the predictions of MaxEntScan obtained with the normal sequence (Figure 25) and the mutant one (Figure 26), an overt difference emerged between the score of the 5´splice produced with the normal sequence, 7.00, which decreases to 3.27 with the mutated one. Once again, this result indicates that the mutation might probably affect the recognition of the 5' motif by the spliceosome and consequently affect the normal splicing pattern of the gene.

4.2.1.4. Human Splicing Finder

Human splicing finder 2.4.1 (HSF) - (http://www.umd.be/HSF/) - is a tool to predict the effects of mutations on splicing signals or to identify splicing motifs in any human sequence (Desmet *et al.*, 2009).

Concerning the sequences under study, HSF produced the matrices shown in figure 27, which reveal that the wild-type 5´splice site was identified with a score of 81.24, while the mutant one had a slightly lower value (79.07). Although this might appear a minor alteration, the program considers as strong sites those presenting consensus values higher than 80 and as less strong sites those with score values ranging between 70 and 80.

BCKDHA c.108+6T>C

Reference sequence **BCKDHA** Gene > **ENST00000269980** Transcript > Exon number: **1** (108 bp)

```
      1 ATGGCGGTAG CGATCGCTGC AGCGAGGGTC TGGCGGCTAA ACCGTGGTTT GAGCCAGGCT GCCCTCCTGC TGCTGCGGCA GCCTGGGGCT CGGGGACTGG
    101 CTAGATCTgt gagtacctgg gccccaggcg gttttcccaa aggg
```

Total sequence length: 144 nucleotides

Mutant sequence

```
      1 ATGGCGGTAG CGATCGCTGC AGCGAGGGTC TGGCGGCTAA ACCGTGGTTT GAGCCAGGCT GCCCTCCTGC TGCTGCGGCA GCCTGGGGCT CGGGGACTGG
    101 CTAGATCTgt gagcacctgg gccccaggcg gttttcccaa aggg
```

Total sequence length: 144 nucleotides

The sequences analyzed in HSF are underlined.

Quick mutation -

Tables - (Show/Hide)

Potential splice sites

HSF Matrices

Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
102	c.102	Acceptor	TAGATCTgtgagta	tagatctgtgagCA	68.87	69.01	NA	+0.2
106	c.106	Donor	TCTgtgagt	TCTgtgagc	81.24	79.07	0	-2.68

Figure 27. Potential splice sites predicted by BDGP program

4.2.1.5. NetGene2

NetGene2 (http://www.cbs.dtu.dk/services/NetGene2/) is another tool that predicts coding region splice sites (Hebsgaard *et al.*, 1996). In the case of the *BCKDHA* sequences, the program has identified the natural splice site in the WT sequence with a confidence of 0.41, while this splice site was no longer identified in the mutated sequence. This result favors again the potential pathogenic effect of the alteration c.108+6T>C, indicating that the mutant 5´splice site is not recognized by the splicing machinery.

4.2.2. Cloning and construction of reporter minigenes

The splicing alteration identified in patient 3 is located in intron 1 of *BCKDHA* gene, a location that raises technical concerns that needed to be overcome viewing the construction of the minigene. Since the first and last exons of a gene lack either the 3´ss or the 5´ss, respectively, it is necessary to construct a rearranged exon that results from the fusion of two exonic parts in such a way that the resulting hybrid exon contains both a 3' and a 5'ss (Berget, 1995) (Gaildrat *et al.,* 2010).

This strategy has been used successfully in several studies from Belén Pérez group in Madrid, who kindly provided the hybrid pSPL3 vector to be used in this study (Arrabal *et al.,* 2011); (Perez *et al.*, 2013).

In the work of Arrabal and collaborators (Arrabal *et al.,* 2011) a hybrid exon containing part of exon 2 (with its 3' ss) fused with exon 1 (with its 5' ss) of *SPR* gene was constructed to study the impact of a mutation located in the last nucleotide of exon 1. More recently, the same hybrid minigene was successfully used and modified to study a mutation in the first exon of *ALDH7A1* (Perez *et al.*, 2013). Given that both studies demonstrated the utility and applicability of the minigene splicing assays to determine the pathogenic impact of nucleotide alterations located in the first exon/intron boundary that might impair the normal splicing pattern of a gene, the strategy was implemented in the present study.

To functionally study the mutation c.108+6T>C in *BCKDHA* gene, our initial approach was to use the minigene sent from Spain (Figure 28) and replace the fragment correspondent to *ALDH7A1* by the fragment of interest amplified from *BCKDHA* gene, maintaining the *SPR* region with its 3´ss. However, the restriction enzyme SalI that needed to be used in this process cuted the pSPL3 vector in two sites, instead of only one as initially expected, therefore impairing the remaining procedure.

To overcome this problem we decided first to construct an intermediate reporter minigene using TOPO vector and then the final splicing minigene vector. These results are presented and discussed below.

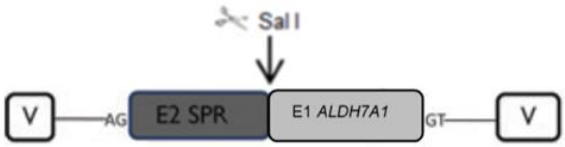


Figure 28. Representation of the pSPL3 splicing vector that contains the hybrid exon formed by part of *SPR* exon 2 fused to *ALDH7A1* exon 1 which was used as final minigene by Perez *et al.* (Perez *et al*., 2013).

4.2.2.1. Construction of an intermediate minigene vector

To construct the intermediate minigene, the region of interest from *BCKDHA* including 203 bp of exon 1 and 301 bp of intron 1, where the mutation is located, was amplified from genomic DNA of a control individual. The region of pSPL3 vector that includes *SPR* fused to *ALDH7A1* was also PCR amplified. As is possible to observe in figure 29, both amplification reactions were successful since strongly stained DNA fragments with the expected sizes were visualized on the agarose gel.

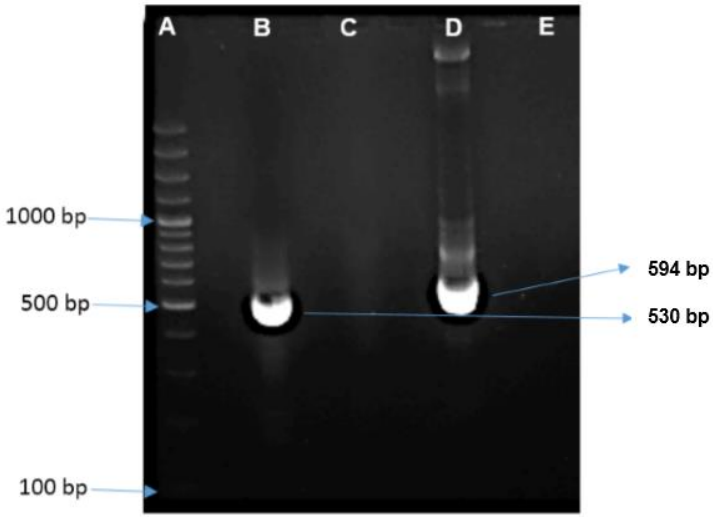


Figure 29. Agarose Gel Electrophoresis of PCR Products **A.** Ladder of 100 bp. **B.** PCR product of fragment from *BCKDHA* gene (530 bp; the size of the fragment is longer than 504bp because of the extra bases included in both primers). **C.** negative control of PCR reaction of fragment from *BCKDHA* gene. **D.** PCR product of fragment from pSPL3 vector (594 bp). **E.** negative control of PCR from fragment of pSPL3 vector.

Both DNA fragments were subsequently digested with SalI enzyme and ligated to construct a hybrid exon formed by *SPR* exon 2 and *BCKDHA* exon 1 – intron 1. This ligation product was then PCR amplified and the efficiency of the reaction was assessed by 1% agarose gel electrophoresis, in which was possible to observe a band of the

expected size (803 bp) (Figure 30). The band was further eluted from the gel and sequenced (Figure 31).

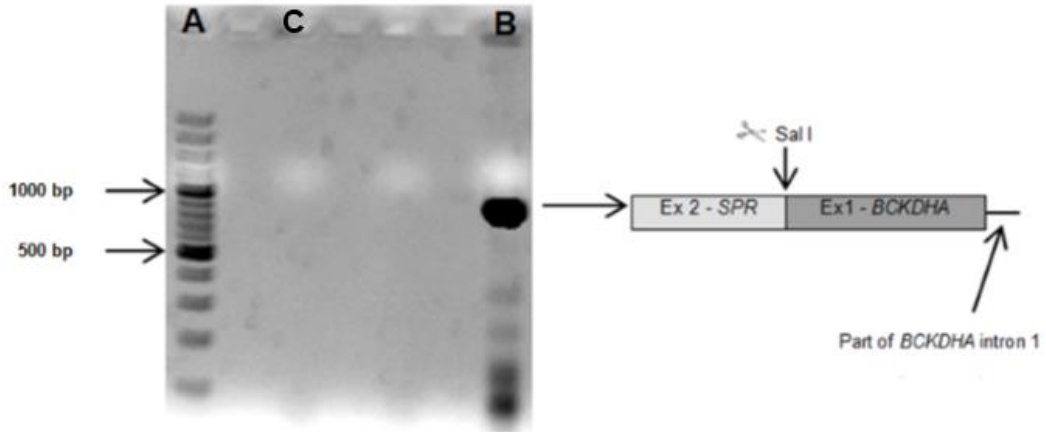


Figure 30. A. Ladder of 1000 bp. **B.** PCR product of the ligation of *SPR* and *BCKDHA* fragments after digestion with restriction enzyme SalI. **C**. Negative control from PCR reaction

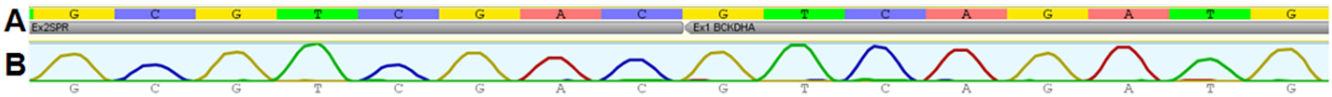


Figure 31. Representation of the ligation between exon 2 from *SPR* gene and exon 1 and part of intron 1 from *BCKDHA* gene **A.** Reference sequence created in Geneious program. **B.** Electropherogram of the fragment corresponding to ligation between *SPR* and *BCKDHA* fragments

This amplified hybrid exon – *SPR+BCKDHA* – was then cloned into pCR®2.1-TOPO vector, and used to transform TOP10 chemically competent E. coli. Colony PCR was implemented to select colonies that integrated the plasmid with the fragment of interest, and the resultant products were then tested in a polyacrylamide gel (Figure 32), in which it was possible to observe that only one colony showed a fragment that could potentially be the one of interest (906 bp). This PCR product was sequenced, which permitted to confirm that it was indeed the expected fragment inserted in the correct orientation.

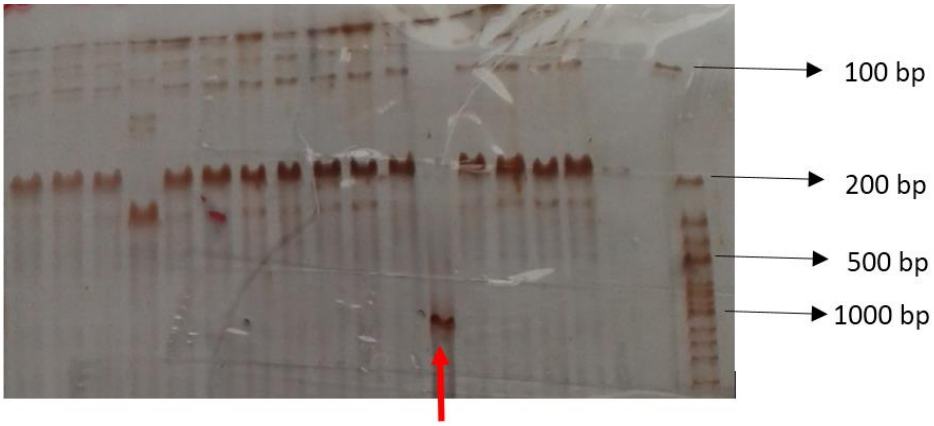


Figure 32. Colony PCR from some of the colonies that grew after the transformation of TOP10 competent cells with TOPO vector with the insert of interest. The colony indicated with the arrow represents the one that has approximately the expected (with 906 bp; PCR performed with TOPO specific primers M13 F and R).

Only one colony had incorporated the insert. This positive colony was maintained in culture using the remaining volume of solution not used in the PCR, while the pipet tip used to pick the colony was dipped into 5 ml of liquid LB medium with ampicillin (50 µg/ml), and was cultured overnight at 37°C to produce numerous copies of intermediate vector.

At the end of this step, the intermediate vector was successfully constructed.

4.2.2.2. Construction of final minigene splicing vector

After the construction of the intermediate minigene vector, created with the purpose of producing numerous copies of the inserted fragment, the final minigene vector was constructed with pSPL3 vector, which was designed to express the protein product coded by the inserted gene.

To construct the final reporter minigene, the intermediate vector that includes the hybrid exon *SPR-BCKDHA* was digested with restriction enzymes BamHI and EcoRV as well as the pSLP3 minigene vector sent from Spain (that includes the hybrid *SPR-ALDH7A1* insert). The products of both digestions were separated by agarose gel electrophoresis and the target fragments were cut from de gel to subsequently elute the DNA.

As can be seen in figure 33, the enzymatic restrictions were well succeed since the distinct fragment were clearly separated from each other and all had the expected sizes. The fragments of interest for the ligation were eluted from the agarose gel, reason that explain not being present in the picture shown in figure 33. These fragments

corresponded to the pSPL3 vector without the “original” insert and to the fragment that includes the hybrid exon *SPR-BCKDHA*, which was ligated to the pSPL3. Once again, the ligation product was used to transform TOP10 competent cells but only one colony grew. Colony PCR from that unique colony was performed to assess the presence and integrity of the insert, which was indeed confirmed since the band of expected size in the agarose gel was eluted and submitted to Sanger sequencing that revealed the presence of the fragment. Finally, DNA was isolated from the culture of that colony and the wild-type/control minigene was successfully constructed.

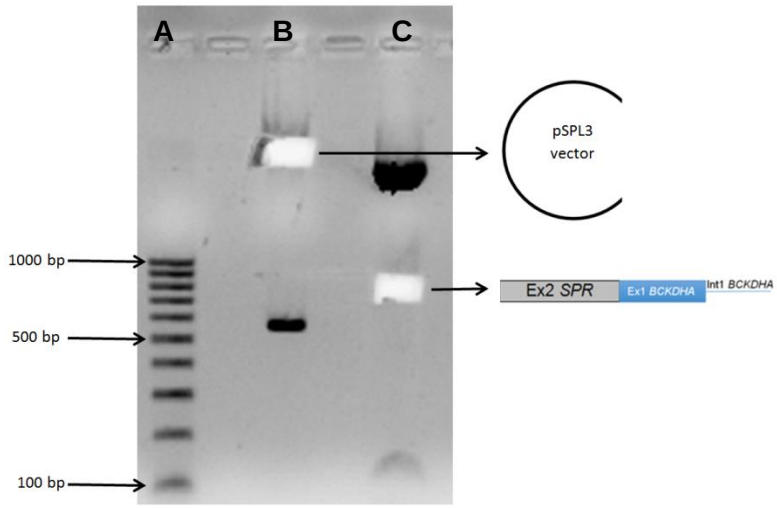


Figure 33. A. Ladder of 100 bp .**B.** Product of enzymatic restriction of pSPL3 vector provided by Belén Pérez with BamHI and EcoRV resulting in a fragment of 5503 bp that correspond to pSPL3 vector without the “original” insert and in a fragment of “original” insert with 528 bp. The band correspondent to pSPL3 vector without the insert was already eluted from the gel when the picture was captured. **C.** Product of enzymatic restriction of intermediate minigene. The results of this reaction are a fragment correspondent to TOPO vector without the insert and the other fragment correspond to the hybrid exon *SPR-BCKDHA* (751 bp) which was excised from the gel.

4.2.2.3. Mutagenesis

As the DNA sample from the Brazilian patient with the mutation c.108+6T>C had not sufficient quality to be used in the minigene construction, we decided to insert the mutation in the previously constructed wild-type reporter splicing minigene, through site-directed mutagenesis. The mutant minigene was successfully constructed since the mutation c.108+6C was introduced and no more alterations were inserted.

In figure 34, it is possible to compare part of the WT minigene sequence with the mutant one.

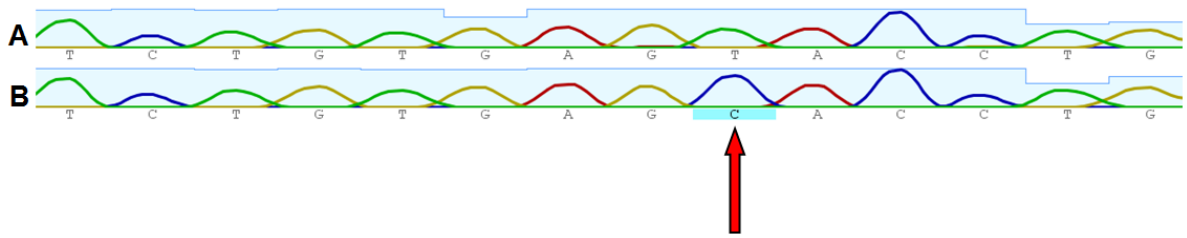


Figure 34. Comparison of the WT sequence with the sequence with the mutation c.108+6T>C to verify the success of the mutagenesis reaction. The red arrow indicates where the mutation was inserted. **A.** WT sequence **B.** sequence with mutation c.108+6T>C.

4.2.2.4. Transfection and cDNA analysis

After obtaining the final minigene with the WT sequence and the minigene with the alteration c.108+6T>C (Figure 35.A), transfection in HeLa and AGS cells was performed with each of these minigenes and, as a control assay, a minigene with the empty vector. These cell lines were selected because they are easy to maintain and transfect, and the comparison of the results obtained in both of them could be useful if ambiguous results were observed (Cooper, 2005). The purpose of these transfections was to verify if the presence of the referred substitution had any effect on mRNA splicing.

After transfection, RNA was extracted from both cell lines, cDNA synthesized, and PCR performed using vector specific primers to amplify transcripts produced in each experiment. Three independent assays were performed reproducing the same conditions to strengthen the quality of results. In all of them, the transcripts produced were concordant as showed the electrophoretic run on 1.6% agarose gels (Figure 35.B).

In the case of the empty pSPL3 used as a control, the size of the band fitted the expectations for an empty vector, which was further confirmed by Sanger sequencing.

Regarding the WT minigene, two transcripts were produced, both of them resulting from correct splicing leading to inclusion of the hybrid exon *SPR-BCKDHA*. The smallest transcript corresponds to fragments in which the splicing occurred at a cryptic 3' splice site within the exonic sequence as previously described by Arrabal and collaborators who have also used a hybrid minigene with *SPR* exon 2. Once again, the integrity of both transcripts was confirmed by Sanger sequencing.

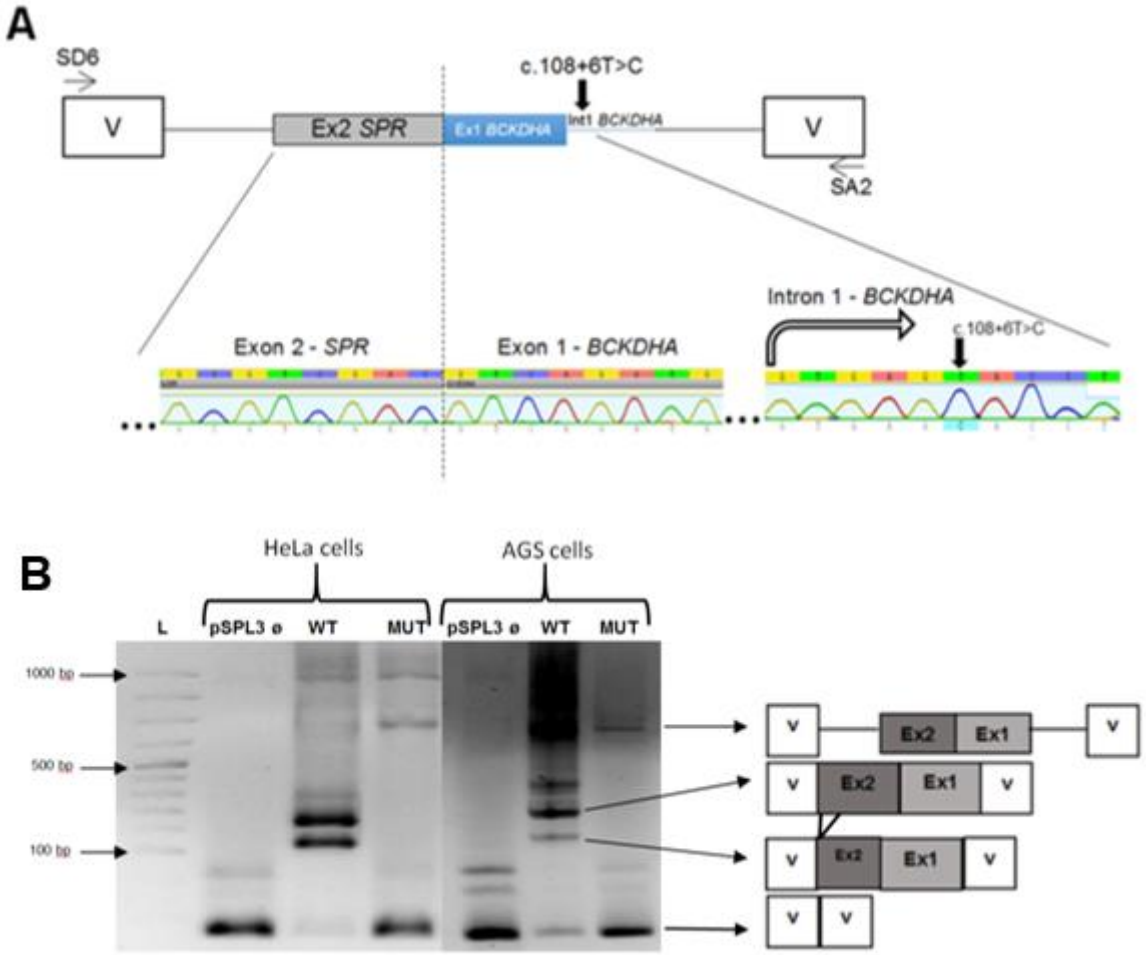


Figure 35. Minigene analysis of the alteration c.108+6T>C. **A.** Schematic drawing of the hybrid exon with the mutation c.108+6T>C cloned in pSPL3 vector with the corresponding sequence analysis. **B.** Post transfection PCR products of the minigene constructs obtained by SD6 and SA2 primers. **L.** Ladder of 100 bp. **pSPL3 ø.** Empty vector. **WT.** WT minigene. **MUT.** Mutant minigene.

In the case of mutant minigene, the transcriptional pattern detected was different from that of the WT (figure 35.B). The major transcript corresponded to the empty vector resulting from splicing between vector sequences, which means that the entire hybrid exon inserted in pSPL3 was skipped. This result indicates that splicing was made with the canonical splice sites of the vector, being therefore plausible to conclude that the substitution c.108+6T>C might cause exon 1 skipping, confirming the suspected disease causative effect of the mutation.

Sequence between -2 and +6 of the splice-donor site are commonly highly conserved (Mount, 1982) and most known mutations involving these sites affect seriously the correct splicing. Although splicing mutations at positions +6 are not very

frequent, some reports already exist describing mutations in similar position to the one where c.108+6T>C lies (+6T>C). For instance, c.313+6T>C in *LDLR* gene, which is associated to familial hypercholesterolaemia (Bourbon *et al.*, 2009), and IVS7+6T>C in *ABCA I* gene that causes low HDL-C (Rhyne *et al.,* 2009). Interestingly, both these mutations cause exon skipping, which to a certain extent mirror our results.

The longer transcript observed in the mutant minigene corresponds to the entire sequence contained in the fragment amplified with vector specific primers, and therefore includes the whole insert introduced in the vector (exon 2 from *SPR* gene, exon 1 and part of intron 1 from *BCKDHA* gene). Probably, this transcript results from a technical artifact of the minigene assay, a secondary result that is known to occur often.

All data here presented clearly indicate that the mutation c.108+6T>C in intron 1 of *BCKDHA* gene impairs the recognition of the 5' splice site by the splicing machinery, being very likely a MSUD causal mutation.

However, because fragments able to be inserted in minigene are of limited size, it was not possible to include the entire *BCKDHA* intron 1 (it has more than 12Kb), and so it cannot be excluded the possibility that a cryptic splice site in the remaining part of intron 1 is used instead of the normal one. Only the analysis of patient´s RNA sample could unequivocally elucidate this uncertainty and clearly elucidate the impact of the mutation.

Nevertheless, as demonstrated in previous studies, the exon-trapping system is a reliable tool that allows to have a prediction of whether a substitution of unknown significance influences splicing *in vivo*, although it is not always possible to fully understand the level of impairment caused.

Despite the limitation, this study affords strong evidence on the pathogenic effect of the alteration c.108+6T>C in homozygosity in *BCKDHA* gene, being very probably the cause of MSUD in patient 3.

Toward that conclusion pointed the results from the bioinformatics analyses undertaken, since all tools indicated a potential alteration in the recognition of the 5' splice site where the substitution is located. Of note the predictions of the programs NNSplice and NetGene2, both indicating that the mutant splice site ceased to be recognized by the splicing machinery, which afterwards was validated by the functional assays performed.

In relation to the molecular spectrum of MSUD, up to now only few splicing alterations have been described, four in *BCKDHA*. Possibly this number corresponds to an underestimate since the screening of intronic regions is not always done when

searching for MSUD causative mutations. This study also illustrates how useful can be the investment in a fine molecular characterization of MSUD patients.

5. Conclusion

This project focused on the molecular characterization of three MSUD patients.

For two of them, the characterization was completely performed in this study, while for the remaining patient it was possible to demonstrate through functional assays the effect of the alteration identified in a previous study.

A total of three mutations were found in *BCKDHB* gene, two of them already known as a cause of MSUD: a deletion of 11 bp (c.93_103del11; p.Ala32Phefs*48) detected in homozygosity in one of the patients, and a nonsense mutation known to be particularly frequent in the Iberia Peninsula, c.799C>T (p.Gln267Ter), identified in heterozygosity in another patient. The third alteration, also detected in heterozygosity, was a missense substitution, c.359T>C (p.Phe120Ser). Since the p.Phe120Ser mutation had never been described before and was not classified as a polymorphism in on-line databases, to infer its putative pathogenic impact several bioinformatics programs were used and a 3D structural analysis was performed, with results consistently indicating the disease causative effect of the mutation.

Considering the third patient, the previous screening of the genes most frequently affected in MSUD lead to the detection of a new alteration in homozygousity, located in intron 1 of *BCKDHA* gene (c.108+6T>C). Given that this substitution was not located in the almost invariant dinucleotides at intron ends, for which pathogenicity is the commonest consequence, and since no RNA sample from the patient was available, alternative tools have been implemented to determine its effect. Specifically, very detailed bioinformatics analyses were conducted to evaluate the impact of the substitution in the recognition of the 5' splice site by the splicing machinery. In addition, a minigene assay was also implemented to evaluate the effects at the mRNA level, once is now widely acknowledged that minigene assays represent an extremely useful and reliable tool to access the impact of substitutions of unknown clinical significance on splicing, with results showing an excellent correlation with those obtained in RNA from cell cultures of patients ((Fernandez-Guerra, Navarrete et al. 2010); (Arrabal, Teresa et al. 2011); (Perez, Gutierrez-Solana et al. 2013)).

The functional characterization of the mutation c.108+6T>C in *BCKDHA* gene here performed involved the successful construction of two hybrid minigenes that contained the exon 1 and part of intron 1 from *BCKDHA* gene with the normal or mutant nucleotide. After transfection in two different cell lines (HeLA and AGS), subsequent RNA

isolation and cDNA synthesis were performed to analyze the transcripts produced by both minigenes. The approach clearly indicated that the alteration c.108+6T>C is pathogenic, probably leading to the skipping of exon 1 of *BCKDHA* gene, a result that otherwise was concordant with the indications provided by the bioinformatics analyses, especially those relying on the NNSplice and NetGene2 programs since both have predicted that the affected 5' splice site was no longer recognized.

Despite the strong evidence obtained pointing to c.108+6T>C as a MSUD causal mutation, since it was not technically possible to include the entire *BCKDHA* intron 1 (it has more than 12 Kb) in the minigene, a cautionary uncertainty persists on whether the skipping of the exon1 is the direct consequence of the alteration. If only the analysis of patient´s RNA sample could unequivocally elucidate this doubt, in this study important steps were taken to understand the impact of this mutations.

6. Bibliography

Ævarsson A, Chuang JL, Wynn RM, Turley S, Chuang DT, Hol WGJ. (2000). "Crystal structure of human branched-chain α-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease". Structure **8** (3): 277–291.

Arrabal, L., L. Teresa, R. Sanchez-Alcudia, M. Castro, C. Medrano, L. Gutierrez-Solana, S. Roldan, A. Ormazabal, C. Perez-Cerda, B. Merinero, B. Perez, R. Artuch, M. Ugarte and L. R. Desviat (2011). "Genotype-phenotype correlations in sepiapterin reductase deficiency. A splicing defect accounts for a new phenotypic variant." Neurogenetics **12**(3): 183-191.

Baralle, D. and Baralle M. (2005). "Splicing in action: assessing disease causing sequence changes." J Med Genet 42(10): 737-748.

Baralle D., Lucassen A. and Buratti E. (2009) “Missed threads. the impact of pre-mRNA splicing defects on clinical practice.” EMBO Rep 2009 Aug;10(8):810-6

Ben-Dov C., Hartmann B., Lundgren J. and Valcárcel J. (2008). "Genome-wide analysis of alternative pre-mRNA splicing." J Biol Chem 283(3): 1229-1233.

Bentley, D. R., S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, J. M. Boutell, J. Bryant, R. J. Carter, R. Keira Cheetham, A. J. Cox, D. J. Ellis, M. R. Flatbush, N. A. Gormley, S. J. Humphray, L. J. Irving, M. S. Karbelashvili, S. M. Kirk, H. Li, X. Liu, K. S. Maisinger, L. J. Murray, B. Obradovic, T. Ost, M. L. Parkinson, M. R. Pratt, I. M. Rasolonjatovo, M. T. Reed, R. Rigatti, C. Rodighiero, M. T. Ross, A. Sabot, S. V. Sankar, A. Scally, G. P. Schroth, M. E. Smith, V. P. Smith, A. Spiridou, P. E. Torrance, S. S. Tzonev, E. H. Vermaas, K. Walter, X. Wu, L. Zhang, M. D. Alam, C. Anastasi, I. C. Aniebo, D. M. Bailey, I. R. Bancarz, S. Banerjee, S. G. Barbour, P. A. Baybayan, V. A. Benoit, K. F. Benson, C. Bevis, P. J. Black, A. Boodhun, J. S. Brennan, J. A. Bridgham, R. C. Brown, A. A. Brown, D. H. Buermann, A. A. Bundu, J. C. Burrows, N. P. Carter, N. Castillo, E. C. M. Chiara, S. Chang, R. Neil Cooley, N. R. Crake, O. O. Dada, K. D. Diakoumakos, B. Dominguez-Fernandez, D. J. Earnshaw, U. C. Egbujor, D. W. Elmore, S. S. Etchin, M. R. Ewan, M. Fedurco, L. J. Fraser, K. V. Fuentes Fajardo, W. Scott Furey, D. George, K. J. Gietzen,

C. P. Goddard, G. S. Golda, P. A. Granieri, D. E. Green, D. L. Gustafson, N. F. Hansen, K. Harnish, C. D. Haudenschild, N. I. Heyer, M. M. Hims, J. T. Ho, A. M. Horgan, K. Hoschler, S. Hurwitz, D. V. Ivanov, M. Q. Johnson, T. James, T. A. Huw Jones, G. D. Kang, T. H. Kerelska, A. D. Kersey, I. Khrebtukova, A. P. Kindwall, Z. Kingsbury, P. I. Kokko-Gonzales, A. Kumar, M. A. Laurent, C. T. Lawley, S. E. Lee, X. Lee, A. K. Liao, J. A. Loch, M. Lok, S. Luo, R. M. Mammen, J. W. Martin, P. G. McCauley, P. McNitt, P. Mehta, K. W. Moon, J. W. Mullens, T. Newington, Z. Ning, B. Ling Ng, S. M. Novo, M. J. O'Neill, M. A. Osborne, A. Osnowski, O. Ostadan, L. L. Paraschos, L. Pickering, A. C. Pike, A. C. Pike, D. Chris Pinkard, D. P. Pliskin, J. Podhasky, V. J. Quijano, C. Raczy, V. H. Rae, S. R. Rawlings, A. Chiva Rodriguez, P. M. Roe, J. Rogers, M. C. Rogert Bacigalupo, N. Romanov, A. Romieu, R. K. Roth, N. J. Rourke, S. T. Ruediger, E. Rusman, R. M. Sanches-Kuiper, M. R. Schenker, J. M. Seoane, R. J. Shaw, M. K. Shiver, S. W. Short, N. L. Sizto, J. P. Sluis, M. A. Smith, J. Ernest Sohna Sohna, E. J. Spence, K. Stevens, N. Sutton, L. Szajkowski, C. L. Tregidgo, G. Turcatti, S. Vandevondele, Y. Verhovsky, S. M. Virk, S. Wakelin, G. C. Walcott, J. Wang, G. J. Worsley, J. Yan, L. Yau, M. Zuerlein, J. Rogers, J. C. Mullikin, M. E. Hurles, N. J. McCooke, J. S. West, F. L. Oaks, P. L. Lundberg, D. Klenerman, R. Durbin and A. J. Smith (2008). "Accurate whole human genome sequencing using reversible terminator chemistry." Nature **456**(7218): 53-59.

Bourbon, M., M. A. Duarte, A. C. Alves, A. M. Medeiros, L. Marques and A. K. Soutar (2009). "Genetic diagnosis of familial hypercholesterolaemia: the importance of functional analysis of potential splice-site mutations." J Med Genet **46**(5): 352-357.

Berry G. T., Heidenreich R., Kaplan P., Levine F., Mazur A., Palmieri M., Yudkoff M. and Segal S. (1991). "Branched-chain amino acid-free parenteral nutrition in the treatment of acute metabolic decompensation in patients with maple syrup urine disease." N Engl J Med **324**(3): 175-179.

Burge C. and Karlin S. (1997). "Prediction of complete gene structures in human genomic DNA." J Mol Biol **268**(1): 78-94.

Burrage, L. C., Nagamani S. C., Campeau P. M. and Lee B. H. (2014). "Branched-chain amino acid metabolism: from rare Mendelian dis-eases to more common disorders." Hum Mol Genet. 2014 Sep 15; 23.

Chuang D. T., Ku L. S. and Cox R. P. (1982). "Biochemical basis of thiamin-responsive maple syrup urine disease." Trans Assoc Am Physicians **95**: 196-204.

Chuang, D. and V. Shih (2001). Maple syrup urine disease (branched-chain ketoaciduria). The Metabolic and Molecular Basis of Inherited Disease. S. CR. New York, McGraw-Hill: 1971-2006.

Chuang, J. L., R. M. Wynn, C. C. Moss, J. L. Song, J. Li, N. Awad, H. Mandel and D. T. Chuang (2004). "Structural and biochemical basis for novel mutations in homozygous Israeli maple syrup urine disease patients: a proposed mechanism for the thiamin-responsive phenotype." J Biol Chem **279**(17): 17792-17800.

Chuang, D. T., J. L. Chuang and R. M. Wynn (2006). "Lessons from genetic disorders of branched-chain amino acid metabolism." J Nutr **136**(1 Suppl): 243S-249S.

Cooper, T. A. (2005). "Use of minigene systems to dissect alternative splicing elements." Methods **37**(4): 331-340.

Damuni, Z., Merryfield M. L., Humphreys J. S. and Reed L. J. (1984). "Purification and properties of branched-chain alpha-keto acid dehydrogenase phosphatase from bovine kidney." Proc Natl Acad Sci U S A **81**(14): 4335-4338.

Dancis, J., Levitz M., Miller S. and Westall R. (1959). "Maple syrup urine disease." Br Med J 1(5114): 91-93.

Dancis, J., Levitz M. and Westall R. G. (1960). "Maple syrup urine disease: branched-chain keto-aciduria." Pediatrics 25: 72-79.

Danner, D. J. and C. B. Doering (1998). "Human mutations affecting branched chain alpha-ketoacid dehydrogenase." Front Biosci **3**: d517-524.

Desmet, F. O., D. Hamroun, M. Lalande, G. Collod-Beroud, M. Claustres and C. Beroud (2009). "Human Splicing Finder: an online bioinformatics tool to predict splicing signals." Nucleic Acids Res **37**(9): e67.

Diaz, V., Camarena C., Vega A., Martínez-Pardo M., Díaz C., López M., Hernández F., Andrés A. and Jara P. (2014). "Liver Transplantation for Classical Maple Syrup Urine Disease: Long-Term Follow-Up." J Pediatr Gastroenterol Nutr.

Douglas, A. G. and M. J. Wood (2011). "RNA splicing: disease and therapy." Brief Funct Genomics 10(3): 151-164.

Edelmann, L., Wasserstein M., Kornreich R., Sansaricq C., Snyderman S. and Diaz G. (2001). "Maple syrup urine disease: identification and carrier-frequency determination of a novel founder mutation in the Ashkenazi Jewish population." Am J Hum Genet 69(4): 863-868.

Eng, L., G. Coutinho, S. Nahas, G. Yeo, R. Tanouye, M. Babaei, T. Dork, C. Burge and R. A. Gatti (2004). "Nonclassical splicing mutations in the coding and noncoding regions of the ATM Gene: maximum entropy estimates of splice junction strengths." Hum Mutat **23**(1): 67-76.

Gaildrat, P., A. Killian, A. Martins, I. Tournier, T. Frebourg and M. Tosi (2010). "Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants." Methods Mol Biol **653**: 249-257.

Eswar, N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper and A. Sali (2006). "Comparative protein structure modeling using Modeller." Curr Protoc Bioinformatics **Chapter 5**: Unit 5 6.

Faiz, A. and J. K. Burgess (2012). "How can microarrays unlock asthma?" J Allergy (Cairo) 2012: 241314.

Faustino, N. A. and T. A. Cooper (2003). "Pre-mRNA splicing and human disease." Genes Dev 17(4): 419-437.

Fekete, G., Plattner R., Crabb D., Zhang B., Harris R., Heerema N. and Palmer C. (1989). "Localization of the human gene for the El alpha subunit of branched chain keto acid dehydrogenase (BCKDHA) to chromosome 19q13.1----q13.2." Cytogenet Cell Genet **50**(4): 236-237.

Fernandez-Guerra, P., R. Navarrete, K. Weisiger, L. R. Desviat, S. Packman, M. Ugarte and P. Rodriguez-Pombo (2010). "Functional characterization of the novel intronic nucleotide change c.288+9C>T within the BCKDHA gene: understanding a variant presentation of maple syrup urine disease." J Inherit Metab Dis **33 Suppl 3**: S191-198.

Gaildrat, P., Killian A., Martins A., Tournier I., Frébourg T. and Tosi M. (2010). "Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants." Methods Mol Biol 653: 249-257.

Gottlieb S. (2003). “The splice of life”. Horizon symposia. Understanding the RNAissance. May 2003.

Harris, R. A., Zhang B., Goodwin G., Kuntz M., Shimomura Y., Rougraff P., Dexter P., Zhao Y., Gibson R. and Crabb D. (1990). "Regulation of the branched-chain alpha-ketoacid dehydrogenase and elucidation of a molecular basis for maple syrup urine disease." Adv Enzyme Regul 30: 245-263.

Harris, R. A., Hawes J., Popov K., Zhao Y., Shimomura Y., Sato J., Jaskiewicz J. and Hurley T. (1997). "Studies on the regulation of the mitochondrial alpha-ketoacid dehydrogenase complexes and their kinases." Adv Enzyme Regul **37**: 271-293.

Harris, R. A., Kobayashi R., Murakami T. and Shimomura Y. (2001). "Regulation of branched-chain alpha-keto acid dehydrogenase kinase expression in rat liver." J Nutr **131**(3): 841S-845S.

Harris, R. A., M. Joshi and N. H. Jeoung (2004). "Mechanisms responsible for regulation of branched-chain amino acid catabolism." Biochem Biophys Res Commun **313**(2): 391-396.

Harris, R. A., Joshi M., Jeoung N. and Obayashi M. (2005). "Overview of the molecular and biochemical basis of branched-chain amino acid catabolism." J Nutr **135**(6 Suppl): 1527S-1530S.

Hebsgaard, S. M., P. G. Korning, N. Tolstrup, J. Engelbrecht, P. Rouze and S. Brunak (1996). "Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information." Nucleic Acids Res **24**(17): 3439-3452.

Human Gene Mutation Database (http://www.hgmd.org/)

Hutson, S. M. and T. R. Hall (1993). "Identification of the mitochondrial branched chain aminotransferase as a branched chain alpha-keto acid transport protein." J Biol Chem **268**(5): 3084-3091.

Hutson, S. M., Sweatt A. and Lanoue K. (2005). "Branched-chain [corrected] amino acid metabolism: implications for establishing safe intakes." J Nutr **135**(6 Suppl): 1557S-1564S.

Kornblihtt, A. R., Schor I., Alló M., Dujardin G., Petrillo E. and Muñoz M. (2013). "Alternative splicing: a pivotal step between eukaryotic tran-scription and translation." Nat Rev Mol Cell Biol **14**(3): 153-165.

Layman, D. K. (2003). "The role of leucine in weight loss diets and glucose homeostasis." J Nutr **133**(1): 261S-267S.

Lewandowska, M. A. (2013). "The missing puzzle piece: splicing mutations." Int J Clin Exp Pathol 6(12): 2675-2682.

Lu, G., Sun H., She P., Youn J., Warburton S., Ping P., Vondriska T., Cai H., Lynch C. and Wang Y. (2009). "Protein phosphatase 2Cm is a critical regulator of branched-chain amino acid catabolism in mice and cultured cells." J Clin Invest 119(6): 1678-1687.

Luco, R. F., Allo M., Schor I., Kornblihtt A. and Misteli T. (2011). "Epigenetics in alternative pre-mRNA splicing." Cell 144(1): 16-26.

Lynch, C. J., Patson B., Anthony J., Vaval A., Jefferson L. and Vary T. (2002). "Leucine is a direct-acting nutrient signal that regulates protein synthesis in adipose tissue." Am J Physiol Endocrinol Metab **283**(3): E503-513.

Marshall L. and DiGeorge A (1981) Maple syrup urine disease in the Old Order Mennonites. Am J Hum Genet Suppl **33**:139A.

Matera, A. G. and Z. Wang (2014). "A day in the life of the spliceosome." Nat Rev Mol Cell Biol **15**(2): 108-121.

Matlin, A. J., Clark F. and Smith C. (2005). "Understanding alternative splicing: towards a cellular code." Nat Rev Mol Cell Biol 6(5): 386-398.

McLaughlin, P. M., Hinshaw J. and Stringer A. (2013). "Maple syrup urine disease (MSUD): a case with long-term follow-up after liver transplantation." Clin Neuropsychol **27**(7): 1199-1217.

Menkes, J. H., Hurst P. Craig J. (1954). "A new syndrome: progressive familial infantile cerebral dysfunc-tion associated with an unusual urinary substance." Pediatrics **14**(5): 462-467.

Menkes, J. H. (1959). "Maple syrup disease; isolation and identification of organic acids in the urine." Pediatrics **23**(2): 348-353.

Mordier, S., Deval C., Béchet D., Tassa A. and Ferrara M. (2000). "Leucine limitation induces autophagy and activation of lysosome-dependent proteolysis in C2C12 myotubes through a mammalian target of rapamycin-independent signaling pathway." J Biol Chem **275**(38): 29900-29906.

Morton, D. H., Strauss K., Robinson D., Puffenberger E. and Kelley R. (2002). "Diagnosis and Treatment of Maple Syrup Disease: A Study of 36 Patients." Pediatrics 109(6): 999-1008.

Mount, S. M. (1982). "A catalogue of splice junction sequences." Nucleic Acids Res **10**(2): 459-472.

Na, G., A. Wolfe, C. Ko, H. Youn, Y. M. Lee, S. J. Byun, I. Jeon and Y. Koo (2013). "A low-copy-number plasmid for retrieval of toxic genes from BACs and generation of conditional targeting constructs." Mol Biotechnol **54**(2): 504-514.

Narayanan, M. P., Menon K. and Vasudevan D. (2013). "Analysis of gene mutations among South Indian patients with maple syrup urine disease: identification of four novel mutations." Indian J Biochem Biophys **50**(5): 442-446.

Nobukuni, Y., H. Mitsubuchi, I. Akaboshi, Y. Indo, F. Endo, A. Yoshioka and I. Matsuda (1991). "Maple syrup urine disease. Complete defect of the E1 beta subunit of the branched chain alpha-ketoacid dehydrogenase complex due to a deletion of an 11-bp repeat sequence which encodes a mitochondrial targeting leader peptide in a family with the disease." J Clin Invest **87**(5): 1862-1866.

Oyarzabal, A., M. Martinez-Pardo, Martínez-Pardo M., Merinero B., Navarrete R., Desviat L., Ugarte M. and Rodríguez-Pombo P. (2013). "A novel regulatory defect in the branched-chain alpha-keto acid dehydrogenase complex due to a mutation in the PPM1K gene causes a mild variant phenotype of maple syrup urine disease." Hum Mutat 34(2): 355-362.

Packman, W., Mehta I., Rafie S., Mehta J., Naldi M. and Mooney K. (2012). "Young adults with MSUD and their transition to adulthood: psy-chosocial issues." J Genet Couns **21**(5): 692-703.

Parrella, T., S. Surrey, A. Iolascon, M. Sartore, R. Heidenreich, G. Diamond, A. Ponzone, O. Guardamagna, A. B. Burlina, R. Cerone and et al. (1994). "Maple syrup urine disease (MSUD): screening for known mutations in Italian patients." J Inherit Metab Dis **17**(6): 652-660.

Perez, B., Gutiérrez-Solana L., Verdú A., Merinero B., Yuste-Checa P., Ruiz-Sala P., Calvo R., Jalan A., Marín L., Campos O., Ruiz M., San Miguel M., Vázquez M., Castro

M., Ferrer I., Navarrete R., Desviat L., Lapunzina P., Ugarte M. and Pérez-Cerdá C. (2013). "Clinical, biochemical, and molecular studies in pyridoxine-dependent epilepsy. Antisense therapy as possible new therapeutic option." Epilepsia 54(2): 239-248.

Popov, K. M., Zhao Y., Shimomura Y., Kuntz M. and Harris R. (1992). "Branched-chain alpha-ketoacid dehydrogenase kinase. Molecular cloning, expression, and sequence similarity with histidine protein kinases." J Biol Chem **267**(19): 13127-13130.

Proud, C. G. (2002). "Regulation of mammalian translation factors by nutrients." Eur J Biochem **269**(22): 5338-5349.

Quental, S., Macedo-Ribeiro S., Matos R., Vilarinho L., Martins E., Teles E., Rodrigues E., Diogo L., Garcia P., Eusébio F., Gaspar A., Sequeira S., Furtado F., Lança I., Amorim A. and Prata M. (2008). "Molecular and structural analyses of maple syrup urine disease and identification of a founder mutation in a Portuguese Gypsy community." Mol Genet Metab 94(2): 148-156.

Quental, S., Gusmão A., Rodríguez-Pombo P., Ugarte M., Vilarinho L., Amorim A. and Prata M. (2009). "Revisiting MSUD in Portuguese Gypsies: evidence for a founder mutation and for a mutational hotspot within the BCKDHA gene." Ann Hum Genet **73**(Pt 3): 298-303.

Reed, L. J., Damuni Z. and Merryfield M. (1985). "Regulation of mammalian pyruvate and branched-chain alpha-keto acid dehydrogenase complexes by phosphorylation-dephosphorylation." Curr Top Cell Regul **27**: 41-49.

Reese, M. G., F. H. Eeckman, D. Kulp and D. Haussler (1997). "Improved splice site detection in Genie." J Comput Biol **4**(3): 311-323.

Robinson, B. H., Taylor J. and Sherwood W. (1977). "Deficiency of dihydrolipoyl dehydrogenase (a component of the pyruvate and alpha-ketoglutarate dehydrogenase complexes): a cause of congenital chronic lactic acidosis in infancy." Pediatr Res 11(12): 1198-1202.

Rodriguez-Pombo, P., R. Navarrete, B. Merinero, P. Gomez-Puertas and M. Ugarte (2006). "Mutational spectrum of maple syrup urine disease in Spain." Hum Mutat **27**(7): 715.

Schadewaldt, P., Bodner-Leidecker A., Hammen H. and Wendel U. (1999). "Significance of L-alloisoleucine in plasma for diagnosis of maple syrup urine disease." Clin Chem 45(10): 1734-1740.

Scherer, S. W., Otulakowski G., Robinson B. and Tsui L. (1991). "Localization of the human dihydrolipoamide dehydrogenase gene (DLD) to 7q31----q32." Cytogenet Cell Genet 56(3-4): 176-177.

Scriver, C., Mackenzie, S., Clow, C. and Delvin, E. (1971). "Thiamine-responsive maple syrup urine disease." Lancet 1: 310-312.

Simon, E., Fingerhut R., Baumkötter J., Konstantopoulou V., Ratschmann R. and Wendel U. (2006). "Maple syrup urine disease: favourable effect of early diagnosis by newborn screening on the neonatal course of the disease." J Inherit Metab Dis **29**(4): 532-537.

Singer, R. H. and M. R. Green (1997). "Compartmentalization of eukaryotic gene expression: causes and effects." Cell **91**(3): 291-294.

Snyderman, S. E., Norton P., Roitman E. and Holt L. (1964). "Maple Syrup Urine Disease, with Particular Reference to Dietotherapy." Pediatrics 34: 454-472.

Strachan T. and Read A. (2010). Human Molecular Genetics, 4th edition

Strauss, K. A., Mazariegos G., Sindhi R., Squires R., Finegold D., Vockley G., Robinson D., Hendrickson C., Virji M., Cropcho L., Puffenberger E., McGhee W., Seward L. and Morton D. (2006). "Elective liver transplantation for the treatment of classical maple syrup urine disease." Am J Transplant 6(3): 557-564.

Strauss, K. A. and D. H. Morton (2003). "Branched-chain ketoacyl dehydrogenase deficiency: Maple Syrup Urine Disease." Curr Treat Options Neurol 5: 329-341.

Strauss, K. A., Wardley B., Robinson D., Hendrickson C., Rider N., Puffenberger E., Shellmer D., Moser A. and Morton D. (2010). "Classical maple syrup urine disease and brain development: prin-ciples of management and formula design." Mol Genet Metab **99**(4): 333-345.

Sun, H., Lu G., Ren S., Chen J. and Wang Y. (2011). "Catabolism of branched-chain amino acids in heart failure: insights from genetic models." Pediatr Cardiol **32**(3): 305-310.

Suryawan, A., Hawes J., Harris R., Shimomura Y., Jenkins A. and Hutson S. (1998). "A molecular model of human branched-chain amino acid metabolism." Am J Clin Nutr 68(1): 72-81.

Tabbouche O., A. S., Harry Mountain (2014). "Identification of three novel mutations by studying the molecular genetics of Maple Syrup Urine Disease (MSUD) in the Lebanese population." Molecular Genetics and Metabolism Reports **1**: 273–279.

Topham, C. M., N. Srinivasan and T. L. Blundell (1997). Prediction of the stability of protein mutants based on structural environment-dependent amino acid substitution and propensity tables.

Williams, J. A., A. E. Carnes and C. P. Hodgson (2009). "Plasmid DNA vaccine vector design: impact on efficacy, safety and upstream production." Biotechnol Adv **27**(4): 353-370.

Wynn, R. M., M. Kato, M. Machius, J. L. Chuang, J. Li, D. R. Tomchick and D. T. Chuang (2004). "Molecular mechanism for regulation of the human mitochondrial branched-chain alpha-ketoacid dehydrogenase complex by phosphorylation." Structure **12**(12): 2185-2196.

Vilarinho, L., Rocha, H., Marcão, A., Sousa, C., Fonseca, H., Bogas, M. and Vaz Osório, R. (2006). "Diagnóstico precoce: resultados preliminares do rastreio metabólico alargado." Acta Pediátrica Portuguesa **37**(5): 186-191.

Vockley, J. and R. Ensenauer (2006). "Isovaleric acidemia: new aspects of genetic and pheno-typic heterogeneity." Am J Med Genet C Semin Med Genet **142C**(2): 95-103.

Ward, A. J. and T. A. Cooper (2010). "The pathobiology of splicing." J Pathol **220**(2): 152-163.

Yeaman, S. J., Bassendine M., Fittes D., Hodgson D., Heseltine L., Brown H., Mutimer D., James O. and Fussey S. (1989). "Regulation of the alpha-keto acid dehydrogenase complexes and their involvement in primary biliary cirrhosis." Ann N Y Acad Sci **573**: 183-191.

Yeo, G. and C. B. Burge (2004). "Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals." J Comput Biol **11**(2-3): 377-394.

Zhang, B., Kuntz M., Goodwin G., Edenberg H., Crabb D. and Harris R. (1989). "cDNA cloning of the E1 alpha subunit of the branched-chain alpha-keto acid dehydrogenase and elucidation of a molecular basis for maple syrup urine disease." Ann N Y Acad Sci 573: 130-136.

Zneimer, S. M., Lau K., Eddy R., Shows T., Chuang J., Chuang D. and Cox R. (1991). "Regional assignment of two genes of the human branched-chain alpha-keto acid dehydrogenase complex: the E1 beta gene (BCKDHB) to chromosome 6p21-22 and the E2 gene (DBT) to chromosome 1p31." Genomics **10**(3): 740-747.